9-2010

Structural Characterization of the Pre-Amyloid Oligomers of β-2-Microglobulin Using Covalent Labeling and Mass Spectrometry

Vanessa Leah Castillo Mendoza
University of Massachusetts Amherst, vanessa.lc.mendoza@gmail.com

STRUCTURAL CHARACTERIZATION OF THE PRE-AMYLOID OLIGOMERS OF β-2-MICROGLOBULIN USING COVALENT LABELING AND MASS SPECTROMETRY

A Dissertation Presented

by

VANESSA LEAH C. MENDOZA

Submitted to the Graduate School of the
University of Massachusetts Amherst in partial fulfillment
of the requirements for the degree of

DOCTOR OF PHILOSOPHY

September 2010

Chemistry

STRUCTURAL CHARACTERIZATION OF THE PRE-AMYLOID OLIGOMERS OF β-2-MICROGLOBULIN USING COVALENT LABELING AND MASS SPECTROMETRY

A Dissertation Presented

by

VANESSA LEAH C. MENDOZA

Approved as to style and content by:

_____
Richard W. Vachet, Chair

_____
Lila M. Gierasch, Member

_____
Igor A. Kaltashov, Member

_____
Stephen J. Eyles, Member

_____
Craig T. Martin, Department Head
Department of Chemistry

DEDICATION

My Family and Joe

ACKNOWLEDGEMENTS

I offer my sincerest gratitude to my mentors, colleagues, friends and families.

My adviser, Richard, who has given me encouragement, guidance and opportunity to grow and learn.

My committee members who have given me helpful suggestions and advice.

My labmates who have made the lab a friendly and cheerful environment.

My Amherst pals who have made Amherst feel like home.

My Boston family, the Gills, who has accepted me into their family and opened their home to me.

My family who has always been my source of strength and inspiration.

My lolo, Joe, who has given me unwavering support, encouragement and love.

ABSTRACT

STRUCTURAL CHARACTERIZATION OF THE PRE-AMYLOID OLIGOEMRS OF β-2-MICROGLOBULIN USING COVALENT LABELING AND MASS SPECTROMETRY

SEPTEMBER 2010

VANESSA LEAH C. MENDOZA

B.Sc., UNIVERSITY OF THE PHILIPPINES DILIMAN

Directed by: Professor Richard W. Vachet

The initial steps involved in the assembly of normally soluble proteins into amyloid fibrils remain unclear, yet over 20 human diseases are associated with proteins that aggregate in this manner. Protein surface modification is a potential means of mapping the interaction sites in early oligomers that precede amyloid formation.

This dissertation focuses on the use of covalent labeling combined with mass spectrometry to elucidate the structural features of Cu(II)-induced β-2-microglobulin (β2m) amyloid formation. An improved covalent modification and MS-based approach for protein surface mapping has been developed to address the need for a reliable approach that ensures protein structural integrity during labeling experiments and provides readily detectable modifications. This approach involves measuring the kinetics of the modification reactions and allows any local perturbations caused by the covalent label to be readily identified and avoided.

This MS-based method has been used to study human β-2-microglobulin (β2m), a monomeric protein that has been shown to aggregate into amyloid fibrils in dialysis patients leading to dialysis-related amyloidosis. Under conditions that lead to β2m amyloid formation, reactions of β2m with three complementary covalent labels have been used to identify the Cu(II) binding site, metal-induced conformational changes, and the oligomeric interfaces. Results confirm that Cu(II) binds to His31 and the N-terminal amine. Binding to these residues causes several structural changes in the N-terminal region and ABED β-sheet which likely enables formation of oligomeric intermediates. The covalent labeling data indicate that the pre-amyloid β2m dimer has an interface that involves the antiparallel arrangement of ABED sheets from two monomers. Moreover, our covalent labeling data allowed us to develop a model for the tetramer in which the interface is mediated by interactions between D strands of one dimer unit and the G strands of another dimer unit.

Lastly, the selective covalent modification approach has been used to delineate the structural changes in β2m after interaction with Cu(II), Ni(II), and Zn(II) and their effect on its aggregation. Our covalent labeling data indicates that the unique effect of Cu(II) appears to be caused by the site at which the metal binds the protein and the conformational changes it induces.

TABLE OF CONTENTS

Page

ACKNOWLEDGEMENTS……...…….v

ABSTRACT……...……vi

LIST OF TABLES……...xiii

LIST OF FIGURES……..……xiv

CHAPTERS

1. PROBING PROTEIN STRUCTURE BY AMINO ACID-SPECIFIC COVALENT LABELING AND MASS SPECTROMETRY…….....1

1.1 Amyloid Fibril Formation……1
1.2 β-2-microglobulin (β2m)…….2
1.3 Protein Structural Information from Mass Spectrometry……6
1.4 Basis of Covalent Labeling with Mass Spectrometry……..7

1.4.1 Analysis and Identification of Labeled Sites…….10
1.4.2 Extent of Modification……...13

1.5 Amino Acid Specific Labels and Examples of Their Applications……...17

1.5.1 Arginine…….18

1.5.1.1 Reaction Mechanism……..18
1.5.1.2 Reaction Conditions……...18
1.5.1.3 Biochemical Problems That Have Been Addressed……..19

1.5.2 Histidine…….21

1.5.2.1 Reaction Mechanism……..21
1.5.2.2 Reaction Conditions……...21
1.5.2.3 Side Reactions……22
1.5.2.4 Other Details of Reactivity……23
1.5.2.5 Biochemical Problems That Have Been Addressed……..23

1.5.3 Lysine…….25

1.5.3.1 Reaction Mechanism……..25
1.5.3.2 Reaction Conditions……...26
1.5.3.3 Biochemical Problems That Have Been Addressed……..26

1.6 Summary…….....29
1.7 References……..……30

2. IMPROVED PROTEIN SURFACE MAPPING USING DIETHYLPYROCARBONATE WITH MASS SPECTROMETRIC DETECTION…….....36

2.1 Introduction……36
2.2 Experimental Procedure…….39

2.2.1 Materials……39
2.2.2 DEPC Modification…….....39
2.2.3 Proteolytic Digestion…….40
2.2.4 HPLC Analysis……..40
2.2.5 Mass Spectrometry…….41
2.2.6 Determination of Solvent Accessibility…….41
2.2.7 Fluorescence Measurements……..42
2.2.8 Circular Dichroism (CD) Experiments……..42

2.3 Results and Discussion……..43

2.3.1 Carbethoxylation of Histidine, Tyrosine, Serine and Threonine…….....43
2.3.2 Dose-Response Curves as Indicators of Changes in Protein Structure…….……56
2.3.3 A Comparison with Structural Information from Spectroscopic Techniques……..62
2.3.4 Concentration Effects…….……66
2.3.5 Factors Affecting the Reactivity of Functional Groups.……67

2.4 Conclusions……...……70
2.5 References……..71

3. COPPER BINDING AND BINDING-INDUCED CONFORMATIONAL CHANGES IN β-2-MICROGLOBULIN……..74

3.1 Introduction……74
3.2 Experimental Procedure…….76

3.2.1 Materials…….....76
3.2.2 β2m Solutions……...……76
3.2.3 Carbethoxylation with DEPC……77
3.2.4 Acetylation with NHSA…….77
3.2.5 BD Modification……77
3.2.6 Proteolytic Digestion…….……77
3.2.7 Instrumentation……..78

3.2.8 Amino Acid Modification Percentage……...79

3.3 Results and Discussion……..82

3.3.1 Covalent Labeling with NHSA……..82
3.3.2 Covalent Labeling with DEPC……...85
3.3.3 Covalent Labeling with Butanedione…….86
3.3.4 Copper Binding to Monomeric β2m……..90

3.4 Conclusions……..98
3.5 References……99

4. STRUCTURE OF THE PRE-AMYLOID DIMER OF β-2-MICROGLOBULIN FROM COVALENT LABELING AND MASS SPECTROMETRY……..103

4.1 Introduction……..103
4.2 Experimental Procedure……..105

4.2.1 Materials……..105
4.2.2 Formation of β2m Oligomers……..105
4.2.3 Covalent Labeling Reactions……...106
4.2.4 Proteolytic Digestion……...106
4.2.5 Desalting……..106
4.2.6 Instrumentation……106
4.2.7 Amino Acid Modification Percentage…….108
4.2.8 Determination of Solvent Accessibility……...108
4.2.9 Docking Calculations……...109

4.3 Results and Discussion……110

4.3.1 Covalent Labeling with NHSA……110
4.3.2 Covalent Labeling with DEPC…….114
4.3.3 Covalent Labeling with Butanedione……...119
4.3.4 Determination of Surface Area of β2m Oligomers……..122
4.3.5 Pre-amyloid β2m Dimer……..125

4.4 Conclusions……..137
4.5 References……138

5. STRUCTURE OF THE PRE-AMYLOID TETRAMER OF β-2-MICROGLOBULIN FROM COVALENT LABELING AND MASS SPECTROMETRY……141

5.1 Introduction……..141
5.2 Experimental Procedure……...142

5.2.1 Materials……..142
5.2.2 Formation of β2m Oligomers……..143
5.2.3 Covalent Labeling Reactions……...143
5.2.4 Proteolytic Digestion……...143
5.2.5 Instrumentation……143
5.2.6 Amino Acid Modification Percentage…….144
5.2.7 Molecular Dynamics Simulations……144

5.3 Results and Discussion……145

5.3.1 Covalent Labeling with NHSA……147
5.3.2 Covalent Labeling with DEPC…….151
5.3.3 Covalent Labeling with Butanedione……...156
5.3.4 Pre-amyloid β2m Tetramer……..159

5.4 Conclusions……..171
5.5 References……171

6. DELINEATING THE METAL-SPECIFIC STRUCTURAL CHANGES IN β-2-MICROGLOBULIN AND THEIR EFFECT ON ITS AGGREGATION…….175

6.1 Introduction……..175
6.2 Experimental Procedure……...176

.

6.2.1 Materials……..176
6.2.2 Formation of β2m Oligomers……..177
6.2.3 Covalent Labeling Reactions……...177
6.2.4 Proteolytic Digestion……...177
6.2.5 Instrumentation……177
6.2.6 Amino Acid Modification Percentage…….178
6.2.7 Circular Dichroism (CD)…….178
6.2.8 Thioflavin T Fluorescence……...178
6.2.9 Size Exclusion Chromatography……..179
6.2.10 Transmission Electron Microscopy (TEM)…….179

6.3 Results and Discussion……180

6.3.1 Effect on β2m Structure……...180
6.3.2 Effect on Oligomerization of β2m……...181
6.3.3 Effect on Amyloid Formation of β2m…….188
6.3.4 Covalent Labeling……189

6.4 Conclusions……..196
6.5 References……196

7. SUMMARY AND FUTURE DIRECTIONS……..199

7.1 Summary……..199
7.2 Future Directions…….203

7.2.1 Improving the Protein Surface Mapping Method……..204

7.2.1.1 Multi-Amino Acid Labeling……204
7.2.1.2 Minimizing Reversibility of Covalent Modifications……..204

7.2.2 Structural Characterization of Cu(II)-Induced Oligomer Assembly of β2m……...205

7.2.3 Structural characterization of Zn(II)-Induced Oligomer Assembly of β2m……...207

7.2.3.1 Zn(II) Binding Site……...207
7.2.3.2 Oligomeric Intermediates…….207

7.3 References……..……..208

BIBLIOGRAPHY……211

LIST OF TABLES

Table Page

2.1 DEPC modification rate coefficients for histidine residues…….45

2.2 DEPC modification rate coefficients for tyrosine, serine and threonine residues…….47

2.3 Peptides used for investigating reversibility of DEPC modification……55

3.1 Summary of the percentage changes for the modified amino acids of β2m in the absence and presence of Cu(II)……..88

4.1 Summary of the percentage changes for the modified amino acids from t=0 to t=2 hours after the addition of Cu(II)……..121

4.2 Proteins used to generate the charge-surface area calibration plot…….123

4.3 Measured surface areas of β2m oligomers…….....125

5.1 Summary of the modification percentages for the modified amino acids at 2 hours (t=2 hours) after the addition of Cu(II) and 2.5 days (t=2.5 days) after the addition of Cu(II)…….....158

6.1 Modification percentages for each modified amino acid in the absence of metal and the presence of Cu(II), Ni(II), and Zn(II)…….190

LIST OF FIGURES

Figure Page

1.1 General mechanism of amyloid formation…….2

1.2 Structure and sequence of β2m. (A) Ribbon diagram of human β2m (PDB 2D4F). (B) Amino acid sequence of β2m showing strand nomenclature……...3

1.3 Proposed model for Cu(II)-induced β2m fibril formation. Monomeric β2m binds Cu(II) and is destabilized. Cu(II)-bound β2m forms a dimer which subsequently forms tetramer I. Tetramer I undergoes a conformational change, releases Cu(II), and undergoes a rearrangement to form tetramer II. Tetramer II eventually leads to the hexamer, presumably via an association with a dimer that releases its bound Cu(II). An oligomer just larger than the hexamer acts as the nucleus that enables fibril or protofibril formation……...5

1.4 Analytical scheme showing covalent labeling with MS detection. The label modifies solvent accessible amino acids. ESI- or MALDI-MS can be used to calculate the extent of modification. The modified protein can then be subjected to proteolysis followed by either MALDI-MS or LC-MS/MS to determine the peptide fragments and/or the specific amino acids that have been modified……...9

1.5 (A). Peptide mass mapping. Labeled peptides are identified from a mass spectrum that shows unmodified and modified fragments. (B) In tandem mass spectra unmodified and modified product ions can be used to identify the specific amino acid that has been modified……12

1.6 Analytical scheme showing how the abundances of the unmodified and modified peptides can identify amino acids that undergo changes in reactivity, which can then help determine regions in a protein that undergo a structural change……15

1.7 Plot of unmodified peptide (or protein) as a function of time (or reagent concentration) showing that a more buried residue His-B (blue) react more slowly and therefore has a shallower slope in the kinetics plot compared to a more exposed residue like His-A (orange)…….17

1.8 Modification reaction of arginine by 2,3-butandeione (BD)……..18

1.9 Reaction showing how borate can stabilize adducts that are formed by arginine with the various dicarbonyl compounds……..19

1.10 Modification reactions of diethylpyrocarbonate (DEPC) with histidine, lysine, tyrosine, serine, and threonine……..22

1.11 Crystal structure of the insulin dimer (PDB code 4INS). The ε-nitrogen (NE) and δ-nitrogen (ND) of His-10 are both accessible. His-5 participates in the insulin core structure and shows an accessible surface for either the ε-nitrogen (NE) or the δ-nitrogen (ND), depending on the monomer……..24

1.12 Modification reactions of lysine residues by sulfo-N-hydroxysuccinimide acetate (NHSA)……26

1.13 Crystal structure of the E9:Im9 complex (PDB code 1EMV). Lys89 and Lys97 are directly involved in the interaction. Lys76 and Lys81 are located in a α-helical region near the interaction surface. Lys55 and Lys63 do not interact with Im9 (orange), and the region of E9 (green) containing these residues possibly undergoes a conformational change upon Im9 binding……29

2.1 Reactions of histidine, tyrosine, threonine, and serine with DEPC……...44

2.2 MS/MS spectra of β2m (A-F), cytochrome C (G-J), and myoglobin (K-Q) showing modification of: (A) Thr4 (fragment Thr4-Tyr10), (B) His13 (fragment Ser11-Lys19), (C) His31 (fragment Val27-Lys41)*, (D) Ser33 (fragment Val27-Lys41)*, (E) His51 (fragment Val49-Trp60), (F) Ser88 (fragment Val82-Lys94), (G) His26 (fragment His26-Arg38), (H) Tyr48 (fragment Thr40-Lys55), (I) Tyr67 (fragment Gly56-Lys72), (J) Tyr74 (fragment Lys73-Lys79), (K) N-terminus or Ser3 (fragment Gly1-Lys16), (L) His36 (fragment Leu32-Lys45), (M) His48 (fragment His48-Lys50), (N) His64 (fragment Lys63-Lys77), (O) His81 (fragment Gly80-Lys96), (P) His97 (fragment His97-Lys102), and (Q) Ser117 (fragment Tyr103-Lys118))…….53

2.3 Changes in DEPC modification percentages over time for peptides containing Ser, His, and Thr residues. Each plot is labeled with the sequence of the peptide that corresponds to the data. The lines are not mathematical fits of the data but are included to help visualize the qualitative trend. The thicker lines are for the peptides containing Thr……55

2.4 Dose-response plots for the reactions of DEPC with (A) cytochrome c and (B) myoglobin. The plot for each protein is produced from the ESI-MS data of the DEPC-treated proteins. The $[P]/[P]_0$ ratio is obtained by dividing the peak area for the unmodified protein by the sum of the peak areas for the modified and unmodified protein. The difference between the [P] and $[P]_o$ is used to determine the concentration of DEPC, [X]. The k values are obtained by dividing the measured slopes by the reaction time……..59

2.5 Expanded view of mass spectrum showing the extent of DEPC modification for the (A) +15 charge state of cytochrome c and (B) +21 charge state of myoglobin. CEt refers to a carbethoxy group, which is added upon reaction with DEPC…….60

2.6 Dose-response plots for selected proteolytic fragments of myoglobin after reactions with DEPC. (A) Proteolytic fragment Leu32-Lys45, which contains His36; (B) Proteolytic fragment His48-Lys56; (C) Proteolytic fragment Lys79-Lys96, which contains His81; (D) Proteolytic fragment Gly1-Lys16 which contains Ser3. The plot for each reactive residue is produced from LC-MS data of the proteolytic digests of the modified protein. The ion abundances for the modified and unmodified peptide fragments containing the residue of interest are used to determine [P], $[P]_o$, [X] and $[X]_o$ in a manner similar to that described in the caption of Figure 2.4……..61

2.7 Fluorescence measurements of (A) cytochrome c and (B) myoglobin after reactions with different concentrations of DEPC. Fluorescence was measured at 340 nm for cytochrome c and 320 nm for myoglobin. The average fluorescence emission wavelength is calculated as described in Section 2.2.7…….64

2.8 Far-UV CD spectra of (A) cytochrome c and (B) myoglobin acquired under native conditions (dotted black line), after a 1 min reaction with a 50-fold excess (5 mM) of DEPC (solid light gray line), in the presence of 1 M urea (dashed black line), and in the presence of 9 M urea (dashed gray line)…….65

2.9 Dose-response plot showing the effect of protein concentration on the DEPC modification rate. The modification rate coefficients, k, are 0.027 ± 0.001 for 10 μM (black straight line) and 0.0268 ± 0.0004 for 100 μM (gray dashed line). The k values are obtained by dividing the measured slope by the reaction time…….67

3.1 Extracted ion chromatograms (EIC)of the unmodified (m/z 623.4) and modified (m/z 644.4 and 665.4) forms of the fragment Ile1-Tyr10 in the absence of Cu(II). Each trace represents one trial. The peaks at ~14 min and ~15 min are the Lys6- and N-terminus-modified fragments, respectively. The peaks at ~18.2 min are the di-modified fragments……..…….81

3.2 (A) Dose-response plot for Lys6 (fragment Thr4-Tyr10) after reaction with NHSA. (B) Dose-response plot for Lys75 (fragment Tyr67-Tyr78) after reaction with NHSA. The plot for each reactive residue is produced from LC-MS data of the proteolytic digests of the modified protein. The $[P]/[P]_o$ ratio is obtained by dividing the peak area for the unmodified fragment by the sum of the peak areas for the modified and unmodified fragments. The difference between the [P] and $[P]_o$ values is used to determine the concentration of NHSA, [X]. (C) Expanded view of mass spectrum showing the extent of acetylation for the +12 charge of β2m at t=0. The second peak of each doublet is due to a fraction of the protein being oxidized at Met99. This residue is easily oxidized during protein storage. The portion of the protein molecules that are oxidized is typically about 20 to 30% of the total protein……..……..84

3.3 (A) Dose-response plot for His13 after reaction with DEPC. (B) Dose-response plot for Tyr26 after reaction with DEPC. The plot for each reactive residue is produced from LC-MS data of the proteolytic digests of the modified protein. The $[P]/[P]_o$ ratio is obtained by dividing the peak area for the unmodified fragment by the sum of the peak areas for the modified and unmodified fragments. The difference between the [P] and $[P]_o$ values is used to determine the concentration of DEPC, [X]. (C) Expanded view of mass spectrum showing the extent of carbethoxylation for the +12 charge of β2m at t=0. The second peak of each doublet is due to a fraction of the protein being oxidized at Met99. This residue is easily oxidized during protein storage. The portion of the protein molecules that are oxidized is typically about 20 to 30% of the total protein……..……...87

3.4 (A) Dose-response plot for β2m after reaction with BD. The plot is produced from the ESI-MS data of the BD-treated proteins. The [P] is the remaining concentration of β2m after reaction with BD. (B) Expanded view of mass spectrum showing the extent of modification for the +12 charge of β2m at t=0. The second peak of each doublet is due to a fraction of the protein being oxidized at Met99. This residue is easily oxidized during protein storage. The portion of the protein molecules that are oxidized is typically about 20 to 30% of the total protein……..…….88

3.5 (A) Ribbon representation of monomeric β2m (PDB ID: 2D4F). Amino acids modified by the covalent labels are shown as sticks. (B) Amino acid sequence of β2m showing strand nomenclature. Black lines show amino acids on each β strand. The internal disulfide bond is shown in blue. The amino acids probed by the covalent labels are shown in red…….…....91

3.6 Dose-response plots for β2m after reaction with (A) DEPC, (B) NHSA, and (C) BD. The plots show that the concentration range in which the protein's global structure is maintained is the same for both native β2m and β2m in the presence of Cu(II)……92

3.7 Structure of β2m showing the Cu(II) binding site at the N-terminal region of the protein. This structure was generated from molecular mechanics calculations of the protein, beginning with the crystal structure of the apo-protein (PDB 1LDS) and after adding Cu(II) to the known binding sites……..93

3.8 Ribbon representation of monomeric β2m (PDB ID 2D4F). Amino acids that decrease (green) and increase (orange) in modification extent upon interaction with Cu(II) are shown as sticks……..94

3.9 A comparison of the crystal structure of monomeric β2m (green, PDB 2D4F) and one of the subunits from the H13F hexamer (blue, PDB 3CIQ) that illustrates the repositioning of Arg3…….96

4.1 Extent of NHSA modification throughout the course of dimer formation. (A) N-terminus. (B) Lys6. (C) Lys19. (D) Lys41. (E) Lys58. (F) Lys75. (G) Ans83. (H) Lys91. (I) Lys94…….114

4.2 Extent of DEPC modification throughout the course of the dimer formation. (A) N-terminus. (B) Th4. (C) Lys6. (D) His13. (E) Lys19. (F) Tyr26. (G) Ser28. (H) His31. (I). Ser33. (J) Lys41. (K) His51. (L) Ser57/Lys58. (M) Tyr63. (N) Tyr67. (O) Lys75. (P) Ser88. (Q) Lys94……118

4.3 Extent of BD modification throughout the course of the dimer formation. (A) Arg3. (B) Arg12. (C) Arg45. (D) Arg97……...120

4.4 (A) Calibration plot obtained from mass spectra of the proteins in Table 4.1, which is used to measure the surface areas of the β2m oligomers. (B) A plot of ion intensity as a function of charge state for the β2m dimer. A Gaussian fit is used to determine the average charge state……..124

4.5 Dimers formed by D-D strand interactions. Amino acids Glu50-Lys58 are shown as spheres. (A) D-D strand interface of the crystallographic dimer of P32A. (B) Interaction of adjacent D strands in one of the dimer units (AB chains) in the H13F hexamer. Inter-strand interactions of the side chains of Glu50 and His51 of one monomer with the side chain of Lys58 and the backbone of Phe56, respectively, of another monomer are shown……..…130

4.6 The probed amino acids that are located on the (A) ABED and (B) CFG sheets of β2m……131

4.7 Two possible dimers formed via ABED-ABED inter-sheet interactions. (A) Anti-parallel ABED-ABED arrangement as seen in the dimer formed by the B and C chains of the H13F mutant (PDB ID: 3CIQ). (B) Parallel ABED-ABED arrangement as calculated by the docking of two energy-minimized Cu-free β2m monomers obtained from the MHC complex (PDB ID: 1DUZ)……...….133

4.8 Interface between the chains B and C in an anti-parallel arrangement in the H13F hexamer (PDB ID: 3CIQ), showing the inter-subunit salt bridges involving Asp59 and Lys19 and the complex salt bridge involving Glu16……...134

5.1 (A) Ribbon representation of monomeric β2m (PDB ID: 2D4F), Cu(II)-bound β2m, and the formation of the dimer by stacking of two antiparallel ABED sheets. Amino acids modified by the covalent labels are shown as green sticks. The amino acids found in the dimer interface (ABED β-sheet) are shown as red sticks. (B) Amino acid sequence of β2m showing strand nomenclature [18]. Black lines show amino acids on each β strand. The internal disulfide bond is shown in blue. The amino acids probed by the covalent labels are shown in green……...146

5.2 Extent of NHSA modification throughout the course of tetramer formation. (A) N-terminus. (B) Lys6. (C) Lys19. (D) Lys41. (E) Lys58. (F) Lys75. (G) Ans83. (H) Lys91. (I) Lys94. The changes in modification during the dimer formation are shown (points before red line) as reference points…….150

5.3 Extent of DEPC modification throughout the course of the tetramer formation. (A) N-terminus. (B) Th4. (C) Lys6. (D) His13. (E) Lys19. (F) Tyr26. (G) Ser28. (H) His31. (I). Ser33. (J) Lys41. (K) His51. (L) Ser57/Lys58. (M) Tyr63. (N) Tyr67. (O) Lys75. (P) Ser88. (Q) Lys94. The changes in modification during the dimer formation are shown (points before red line) as reference points……..155

5.4 Extent of BD modification throughout the course of the tetramer formation. (A) Arg3. (B) Arg12. (C) Arg45. (D) Arg97. The changes in modification during the dimer formation are shown (points before red line) as reference points……..157

5.5 Possible tetramer formed by two dimer units with anti-parallel ABED interface as seen in the H13F hexamer [9]. The side chains involved in the interface are shown as spheres. (A) Possible tetramer formed by D strands and BC loops from two dimer units in the crystal structure of H13F hexamer. One dimer unit (chains BC) is shown in gray and the other (chains DE) is shown in green. Chains C and D, which form tetramer interface, are shown in dark gray and dark green, respectively. (B) Interaction of adjacent D strands. Inter-strand interactions of the side chains of Glu50 and His51 of one monomer with the side chain of Lys58 and the backbone of Phe56, respectively, of another monomer are shown. These interactions make Ser57 less exposed to solvent. (C) Six of the amino acids that decrease in reactivity upon tetramer formation, namely, Thr4, Lys41, Arg45, Lys91, Lys94, and Arg97, are solvent accessible in the H13F structure……165

5.6 Interface formed by D-D strand interactions in P32A. (A) Crystallographic dimer formed by P32A. (B) Amino acids Glu50-Lys58 on the D strands are shown as spheres. Inter-strand interactions of the side chains of Glu50 and His51 of one monomer with the side chain of Lys58 and the backbone of Phe56, respectively, of another monomer are shown……167

5.7 The probed amino acids that decrease in modification upon tetramer formation (shown as red sticks) are located on the edge strands, A, C, D and G……..169

5.8 Proposed model structure for the wild-type β2m tetramer. (A) Tetramer formed by the interaction of two D strands from one dimer unit (dark and light gray) and two G strands of another dimer unit (dark and light). (B) Top-view of the tetramer model. (C) Amino acids Thr4 on the N-terminal strand, Glu50-Lys58 on the D strand, and Gln89-Arg97 on the G strand are shown as spheres. Several interactions are shown: salt bridge between Glu50 and Arg97, salt bridge between His51 and Asp96, complex salt bridge between Asp53 and Lys94 and Lys91, and H-bonding between Gln89 and Lys58. (D) These interactions alter the orientation of the D strand such that Arg45 can form a salt bridge with Glu47……170

6.1 CD spectra of β2m in the absence (black) and presence of divalent metal ions: Cu(II) (blue), Ni(II) (green), and Zn(II) (orange). (A) Far-UV CD spectra. (B) Near-UV CD spectra……..181

6.2 β2m oligomerization monitored by thioflavin T (ThT) fluorescence at 483 nm. The ThT fluorescence maximum shifts from 450 nm to 483 nm upon binding to β2m oligomers or amyloid-like species in solution. The trend lines associated with the plots of the metal--containing samples are from non-linear curve fitting of the data……..183

6.3 β2m oligomerization in the presence of Ni(II) monitored by SEC. The monomer peak is observed at ~27 min…….184

6.4 SEC and SEC-MS analyses of β2m incubated with Cu(II). (A) Chromatograms for a control (black) and samples incubated with Cu(II) for 1 day (red), 3 days (blue), and 6 days (green). Monomer (M), dimer (M_2), tetramer (M_4), and hexamer (M_6) are observed. (B) Temporal progression of the oligomers measured by SEC. (C) SEC-ESI mass spectrum of the chromatographic peak eluting at ~25 min after 6 days of incubation, confirming that this peak corresponds to the β2m dimer. The charge state of each mass spectral peak is labeled……..185

6.5 SEC and SEC-MS analyses of β2m incubated with Zn(II). (A) Chromatograms for a control (black) and samples incubated for 1 day (red), 3 days ((blue), and 6 days (green). Monomer (M), dimer (M_2), oligomers bigger than a hexamer are observed. (B) Temporal progression of the oligomers measured by SEC. (C) SEC-ESI mass spectrum of the chromatographic peak eluting at ~23 min after 6 days incubation, confirming that this is the β2m dimer. The charge state of each mass spectral peak is labeled……187

6.6 (A) Transmission electron micrograph images obtained from β2m after incubation with Zn(II) for 1 month. Dimensions of image is 200 nm x 325 nm. (B) The TEM image for β2m incubated with Cu(II) for 1 month is adapted from reference 12……189

6.7 Ribbon representation of monomeric β2m (PDB ID: 2D4F). Amino acids that decrease (green) and increase (orange) in reactivity upon interaction with divalent metal ions are shown as sticks. (A) Cu(II). (B) Ni(II). (C) Zn(II)……195

CHAPTER 1

PROBING PROTEIN STRUCTURE BY AMINO ACID-SPECIFIC COVALENT LABELING AND MASS SPECTROMETRY

This chapter is part of a paper published as: Mendoza, V. L., and Vachet, R. W. (2009) Probing protein structure by amino acid-specific covalent labeling and mass spectrometry. *Mass Spec. Rev. 28*, 785-815.

1.1 Amyloid Fibril Formation

A number of proteins are known to aggregate into amyloid fibrils *in vivo*, leading to a pathological disorder known as amyloidosis. Amyloidoses are protein conformational diseases that arise from the self-assembly of normally soluble proteins into insoluble fibrils, which accumulate in a variety of tissues and organs [1-4]. One of the striking characteristics of these pathological states is that despite the variability in the amino acid sequences and tertiary structures of the precursor proteins, the fibrils are very similar in their overall appearance, morphology and properties. X-ray diffraction analyses indicate that all amyloid fibrils share a common cross-β structure, wherein the β strands are arranged orthogonally to the fibril axis and hydrogen bonding parallel to the fibril axis [5,6]. Electron microscopy studies reveal that the fibrils are usually long, straight, and unbranched and have repeating units that reflect the twisting of the component filaments around one another [7]. Typically, amyloid fibrils are identified by their green birefringence and ability to bind the dyes Congo Red and thioflavin T.

Another common feature exhibited by these amyloid systems is the presence of a lag phase, which is indicative of nucleation-dependent kinetics [8]. These common characteristics exhibited by the amyloid formation from different precursor proteins suggest a general mechanism for amyloid formation (Figure 1.1) [8]. The process begins

with the formation of a partially folded intermediate. The partially folded conformers self-associate into a nucleus enriched with β sheets. β sheet extension then leads to formation of protofibrils, which eventually give rise to mature fibrils.

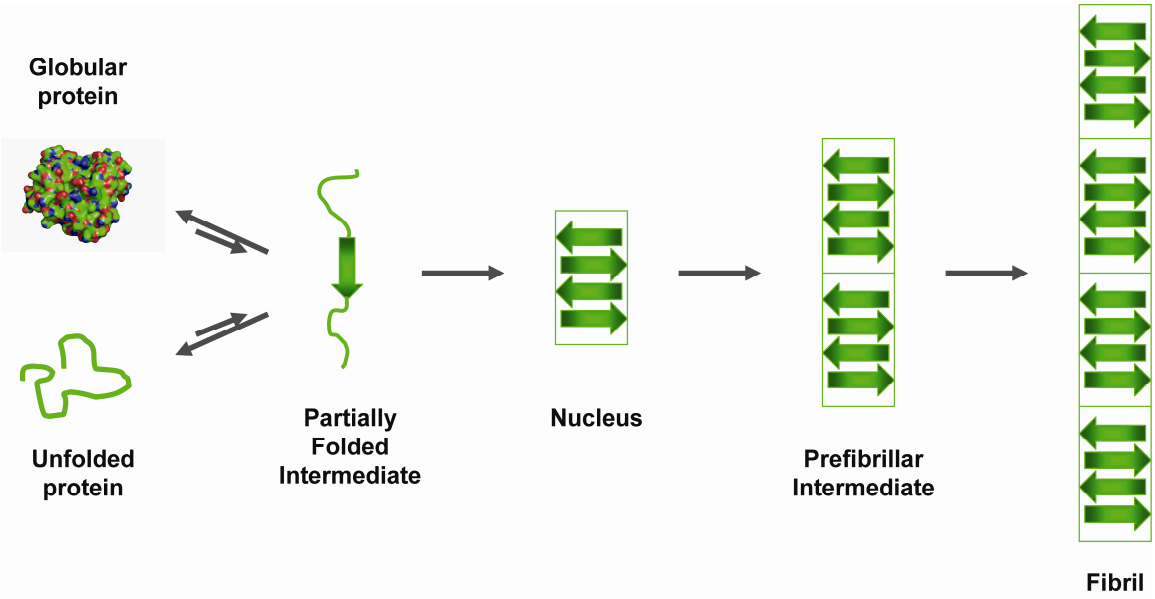


Figure 1.1 General mechanism of amyloid formation.

1.2 β-2-microglobulin (β2m)

One protein that forms amyloid fibrils is β-2 microglobulin (β2m). β2m can accumulate as amyloid fibrils in the musculoskeletal system as a complication of long-term hemodialysis, leading to a condition known as dialysis-related amyloidosis (DRA) [9]. β2m is the non-covalently bound light chain of the class I major histocompatibility complex (MHC-I). It has 99 residues (~12 kDa) and adopts the immunoglobulin fold with seven β strands forming a β-sandwich in its native state (Figure 1.2) [10]. One β sheet is formed by strands A, B, D and E and the other consists of strands C, F and G. A disulfide bond between Cys25 and Cys80 tethers strands B and F in the folded state of the protein.

β2m is vital for the correct folding and assembly of the major histocompatibility complex I (MHC-I) complex that is required for antigen display [10]. As part of normal cell turnover, β2m is released from MHC and carried to the kidney where it is degraded. Upon renal failure, serum levels of β2m increase up to ~60 times above the normal level of 0.1 μM, and the protein aggregates into insoluble amyloid deposits. However, elevated β2m levels are not unique to renal failure patients and are not sufficient to trigger fibrillogenesis [11,12]. β2m amyloid formation must therefore result from features unique to hemodialysis.

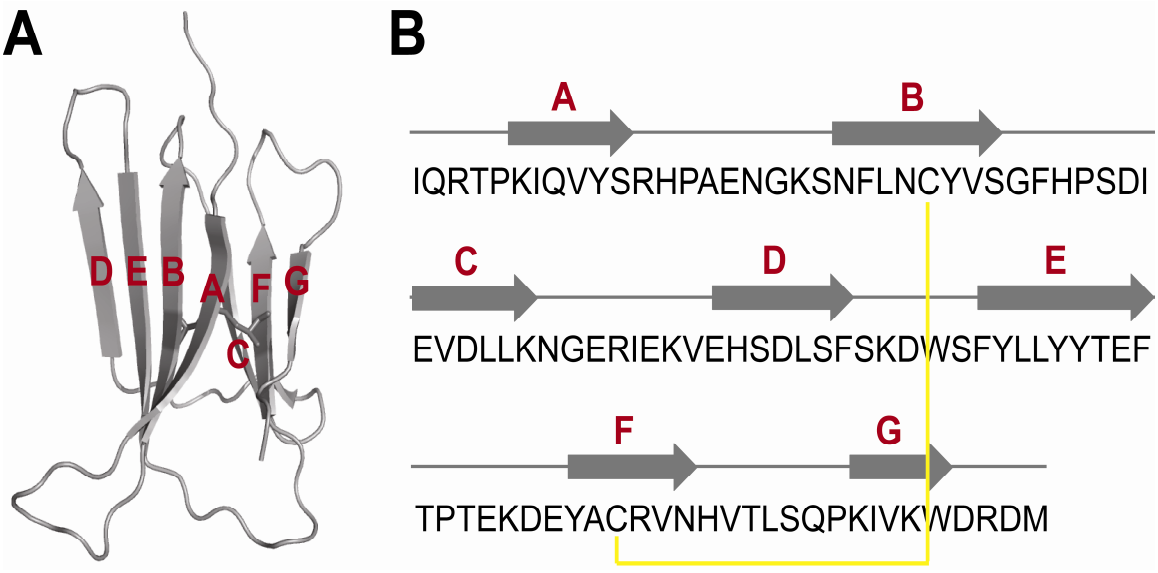


Figure 1.2 Structure and sequence of β2m. (A) Ribbon diagram of human β2m (PDB 2D4F). (B) Amino acid sequence of β2m showing strand nomenclature [13].

Dialysis patients are exposed to both elevated levels of β2m and transition metal cations, particularly Cu(II) [14]. The maximum level of Cu(II) allowed by the FDA in the dialysate is 1.6 μM, which is close to the measured dissociation constant for the β2m

complex with Cu(II) [15]. Also, some dialysis membranes typically contain ~2 mg Cu/m^2 and, hence, provide high concentrations of Cu(II) [15]. Biophysical studies have shown that *in vitro* interactions of stoichiometric amounts of β2m and Cu(II) under near physiological conditions result in fibril formation [14-17]. This observation suggests that Cu(II) interactions induce conformational changes in β2m and increase its amyloidogenic propensity.

Cu(II)-induced β2m amyloid formation is preceded by the formation of discrete, oligomeric intermediates (Figure 1.3) [16-19]. Characterizing reaction intermediates is essential for understanding any biochemical reaction more deeply, and this is no less true for amyloid fibril forming reactions. The formation of discrete oligomers as prefibrillar intermediates has been observed for several other amyloidogenic proteins too [20-22], and indeed acid-induced (pH = 2.5) amyloid formation of β2m has been found to be preceded by discrete oligomers [23]. Delineating the mechanism of the early stages of β2m fibrillogenesis and the structural properties of these oligomeric intermediates may be critical for developing therapeutic strategies against DRA. Furthermore, recent studies of other amyloid systems suggest that prefibrillar intermediates might be responsible for cellular toxicity rather than the amyloid fibrils [24,25]. Consequently, any methods or results that give insight into intermediates that precede β2m fibrils may be useful for understanding other amyloid systems as well.

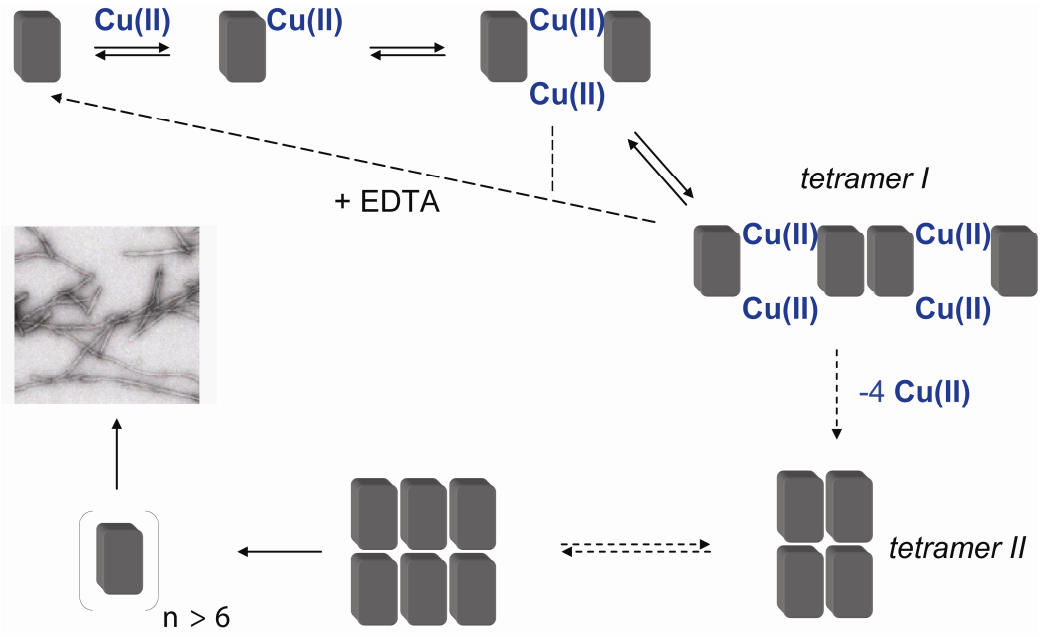


Figure 1.3 Proposed model for Cu(II)-induced β2m fibril formation [17]. Monomeric β2m binds Cu(II) and is destabilized. Cu(II)-bound β2m forms a dimer which subsequently forms tetramer I. Tetramer I undergoes a conformational change, releases Cu(II), and undergoes a rearrangement to form tetramer II. Tetramer II eventually leads to the hexamer, presumably via an association with a dimer that releases its bound Cu(II). An oligomer just larger than the hexamer acts as the nucleus that enables fibril or protofibril formation.

Structural and kinetic studies of the intermediates of the Cu(II)-induced amyloid formation reaction are challenging because these oligomers are in equilibrium, they are transient, and they are present as a mixture. The structural properties of β2m at the onset of fibrillogenesis have been studied using techniques such as circular dichroism (CD), fluorescence, X-ray crystallography, and nuclear magnetic resonance (NMR) [14,26-28]. The morphologies of β2m amyloid fibrils have been elucidated by electron microscopy (EM) and atomic force microscopy (AFM), and some information into how the protein packs in the amyloid fibrils has been obtained by H/D exchange and NMR spectroscopy [27,29,30].

While these studies provide useful insight into β2m fibrillogenesis, the structure of the oligomeric intermediates and the details of the amyloid formation mechanism remain mostly unclear, partly because of limitations of the techniques used in previous studies. CD, for example, gives a measure of the protein's secondary structure but does not provide amino-level information. Fluorescence is very sensitive to conformational changes as well as intermolecular interactions but measurements can be interpreted in structural terms only if the macromolecular structure is known. NMR and X-ray crystallography are the two most important techniques for obtaining high resolution structural information. X-ray crystallography provides detailed structural information but only for the most stable protein state. NMR spectroscopy is restricted by the relatively high sample concentrations required for analysis (mM range), a limited molecular weight range (~30 kDa), difficulties with the analysis of mixtures, and interferences from paramagnetic metals such as Cu(II). Because of the technical limitations of these techniques, mass spectrometry (MS) is playing an ever-increasing role in protein structure determination due to its speed, sensitivity, and specificity. This dissertation focuses on the development and application of MS-based tools to study the structure features of the oligomers that precede β2m fibril formation.

1.3 Protein Structural Information from Mass Spectrometry

With mass spectrometry, structural information about proteins in solution is typically obtained by changing the mass of the protein or its proteolytic fragments in a structure-dependent manner. This indirect way is necessary because MS measurements occur in the gas-phase. Several approaches have been used to achieve this, including H/D

exchange, intra- and intermolecular cross linking, and protein surface mapping via covalent labeling. Covalent labeling entails the use of reagents that either modify specific amino acids or react generally with many different amino acids. In contrast to H/D exchange, the possibilities of back exchange and scrambling are essentially non-existent with covalent labeling reagents. In addition, covalent labeling techniques provide structural information about amino acid side chains, which can be complementary to HDX methods. However, protein structure is more likely to be perturbed with these labels due to the size of typical covalent reagents.

Covalent labeling is generally carried out using either non-specific labels or amino acid-specific labels. Non-specific covalent labeling approaches almost exclusively use protein reactions with radicals (e.g., oxidative footprinting methods). Amino acid-specific labels have been used quite extensively in conjunction with MS to study protein structure and interactions. The sections below will describe particular amino acid-specific covalent labeling reagents and techniques that have been used with MS that can be used to study Cu(II)-induced β2m oligomer assembly under near physiological conditions. The reagents that are used for amino acid-specific labeling, cautionary notes for the proper utilization of these labels, and the protein structural problems that have been addressed with these labels are discussed.

1.4 Basis of Covalent Labeling with Mass Spectrometry

The protein structural information obtained using covalent labeling methods typically rely on the differential reactivity of amino acids upon exposure to a particular labeling reagent. In these experiments the assumption is that amino acids that are exposed

to solvent and therefore accessible to a labeling reagent will be modified, whereas buried amino acids will be modified slowly or not at all. Protein conformational changes, ligand binding, or oligomerization will then affect the extent to which certain amino acid(s) react with the added labeling reagent (Figure 1.4). The amino acids that become more or less accessible to the reagent will react to greater or lesser extents, and amino acid-specific information about a protein's structural changes is then inferred from their differential modification patterns. Covalent labeling has been used for at least 40 years to map protein structure, but the more recent ability of MS to quickly, sensitively, accurately, and precisely map protein modification sites has made covalent labeling approaches much more powerful methods for obtaining protein structural information.

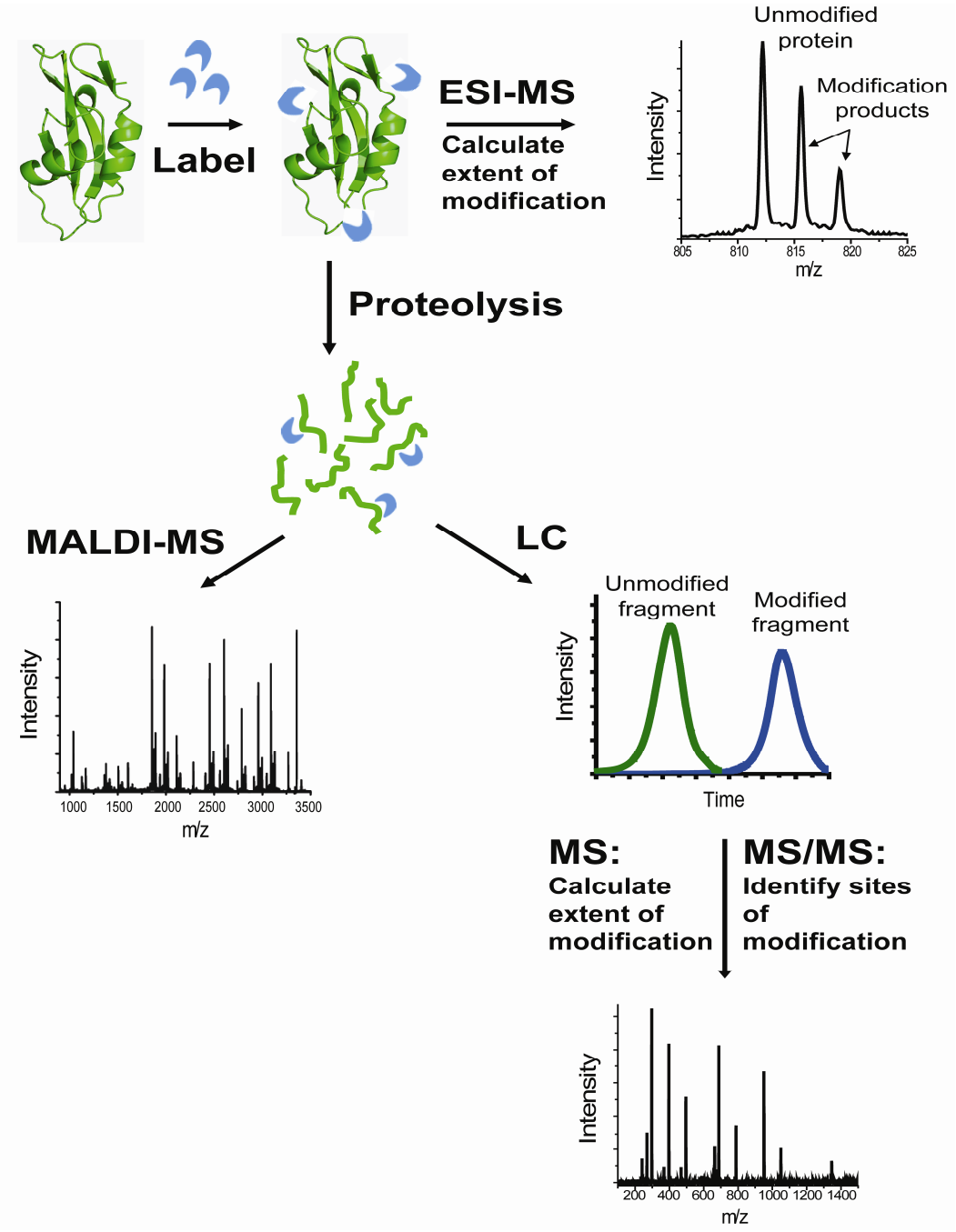


Figure 1.4 Analytical scheme showing covalent labeling with MS detection. The label modifies solvent accessible amino acids. ESI- or MALDI-MS can be used to calculate the extent of modification. The modified protein can then be subjected to proteolysis followed by either MALDI-MS or LC-MS/MS to determine the peptide fragments and/or the specific amino acids that have been modified.

1.4.1 Analysis and Identification of Labeled Sites

The analysis of covalent modifications by MS can be carried out in several ways. The simplest but perhaps least informative means is to determine the number of reagent adducts that modify the protein of interest. The maximum number of labels and/or the average number of labels, which can provide "low resolution" information, can be determined from the mass spectrum of the intact modified protein. This approach is most useful when comparing two protein states and when using a highly specific amino-acid labeling reagent. A change in the amount of modification in one state as compared to another can indicate that a specific amino acid is involved in a structural change or is near a ligand binding site. One example is the study by Leitner and Lindner wherein a combination of 2,3-butanedione and phenylboronic acid was used to modify Arg residues in myoglobin with apomyoglobin (i.e., no heme) [31]. Electrospray ionization (ESI)-MS of the intact proteins indicated that the average number of modified Arg residues increased in apomyoglobin, which is consistent with one or more of these residues being partially buried next to the heme group in the holoprotein.

While monitoring the modification extent of a whole protein by MS is quick and straightforward, it does not provide the same level of detail (or resolution) that is possible when modified proteins are enzymatically digested and the resulting peptide fragments analyzed by ESI-MS or matrix-assisted laser desorption/ionization (MALDI)-MS. Thus, the most common approaches for using covalent modifications to derive protein structural information are peptide mass mapping using MALDI-MS or ESI-MS and peptide sequencing using tandem MS (MS/MS). With peptide mass mapping (i.e. no MS/MS), the modified amino acids are identified by measuring the m/z ratios of the

proteolytic fragments and finding the fragments whose m/z ratios differ, by the mass of label, from the predicted values (Figure 1.5). If an amino-acid specific label is used, then the specific amino acid modified in a proteolytic fragment can usually be inferred. On the other hand, if the relative specificity of the reagent is low, then the identity of the modified amino acid often cannot be assigned with complete confidence. Nonetheless, localizing a modification to a particular proteolytic fragment still improves spatial resolution so that more confident conclusions about a protein's structure can be made. Tandem MS is usually required to identify the specific amino acids modified in a proteolytic fragment. In effect, MS/MS can increase the resolution of the structural information to the single amino-acid level. Obtaining a complete peptide sequence by MS/MS, however, is not always possible, but typically modified amino acids can be at least narrowed down to within 2 to 3 residues.

When the extent of protein modification is low or a mixture of proteolytic peptide fragments to be analyzed is complex, identifying specific modification sites can be challenging. A powerful approach to address this problem is to use modification specific product ions as markers of the modified peptides. The best example of this idea is found in the context of lysine acetylation experiments. Dissociation of peptides with acetylated lysine residues often results in the formation of abundant immonium ions of acetylated lysine at m/z 143, and if extracted ion chromatograms of m/z 143 are used, peptides with modified lysine residues can be readily identified. Another common product ion of lysine-acetylated peptides at m/z 126 has also been used to facilitate the identification of peptides with aceylated lysines [32]. Another means of identifying modified peptides in a mixture of peptides is to use multiple, equally reactive labels that differ in mass.

Recently, Gabant et al. demonstrated that three lysine specific reagents that differed in mass by 113 Da could be used in parallel [33]. The reactivity of these reagents with the protein of interest is virtually identical, causing modified peptides to show up as a quartet of peaks. Finding the modified peptides was facilitated by identifying peptide fragments that had corresponding peaks at m/z values of M+226, M+339, and M+452.

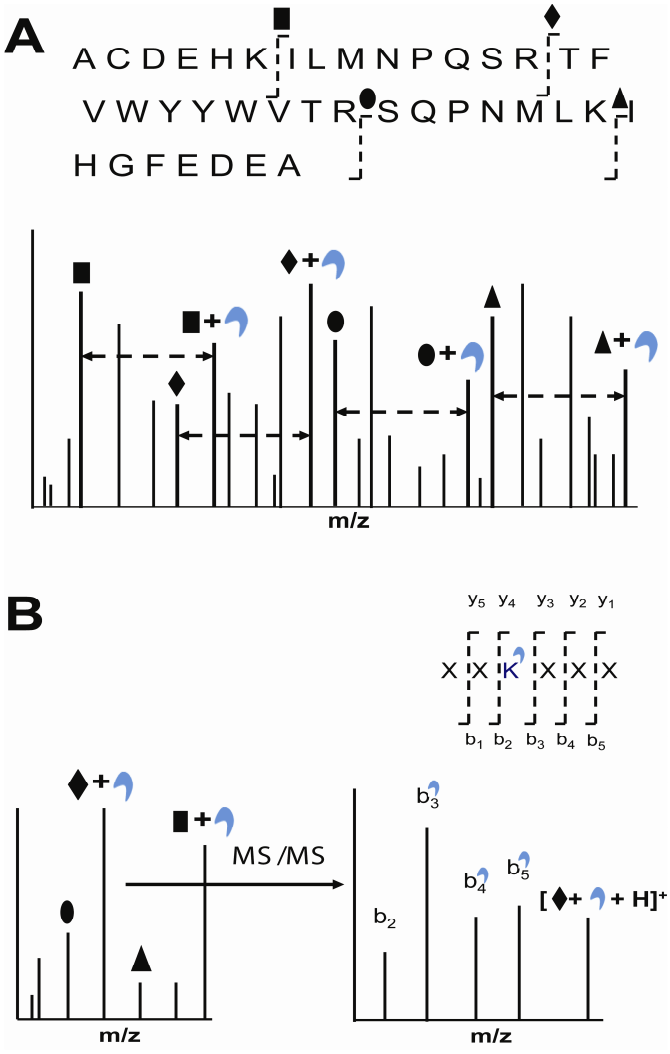


Figure 1.5 (A). Peptide mass mapping. Labeled peptides are identified from a mass spectrum that shows unmodified and modified fragments. (B) In tandem mass spectra unmodified and modified product ions can be used to identify the specific amino acid that has been modified.

An alternative to enzymatic digestion followed by peptide mass mapping or MS/MS is top-down sequencing, which entails performing MS/MS on the whole protein. One of the advantages of the top-down approach is that the protein sequence and any modification site(s) can be determined potentially in a single MS/MS experiment. In addition, the enzymatic digestion step can be avoided. This allows for faster analysis and avoids loss of covalent labels that might not be stable during the relatively long digestion step. However, a large number of product ions are generated by the protein dissociation that can complicate spectral interpretation. Despite this challenge, top-down sequencing, especially on Fourier transform ion cyclotron resonance (FT-ICR) mass spectrometers, is finding more widespread usage [34,35]. Using top-down sequencing to find covalent modification sites, however, presents additional challenges. Multiple potential sites of modification can lead to mixtures of protein isoforms. Upon dissociation, these isobaric protein ions generate even more complicated product ion spectra. Furthermore, for labels with relatively low specificity, narrowing down the modification site to individual amino acids can be difficult. Nonetheless, top-down sequencing has been used effectively for mapping labeled sites so that protein structural information can be obtained [36].

1.4.2 Extent of Modification

Aside from the identification of the modified amino acids, which helps identify the region(s) in a protein that undergoes a structural change or forms an interface with another molecule, the degree of modification at this amino acid is also important to determine. The degree of modification usually reflects the solvent accessibility of the amino acid side chain. Generally, changes in the extent of modification are used as

indicators of the extent of a protein's local structural change. For example, the reactivity of an amino acid that mediates interactions in an interface with another molecule is likely to decrease more than the reactivity of an amino acid that is at the periphery of the interface. Given that the degree of modification can provide plenty of useful information, several approaches have been developed and used to measure amino acid reactivities with modification reagents. The most common approach is to simply compare the extent of modification of the peptide fragment that contains the amino acid under two different conditions – one condition in which the protein has its native structure and another in which the protein undergoes the structural change of interest (Figure 1.6). The ratio of the modified peptide fragment's ion abundance to the peptide fragment's total ion abundance (i.e. unmodified + modified) allows one to easily identify the degree to which an amino acid undergoes a reactivity change that coincides with the protein's structural change. While this approach is straightforward, it does have the potential to give misleading information. Some amino acid modifications can change the ionization efficiency of the measured peptide so that the ion abundance ratio of the modified peptide does not accurately reflect its solution concentration. This problem is likely to be exacerbated if LC is used because modified and unmodified peptides will usually be separated from one another and thus be electrosprayed under different solvent conditions if solvent gradients are used. Furthermore, modification of residues such as lysine that are proteolytic cleavage sites could cause proteolytic enzymes such as trypsin to miss cleavage sites in modified copies of a protein but not in the unmodified copies of the protein. Such an occurrence could also lead to misleading ion abundance ratios if not properly accounted for.

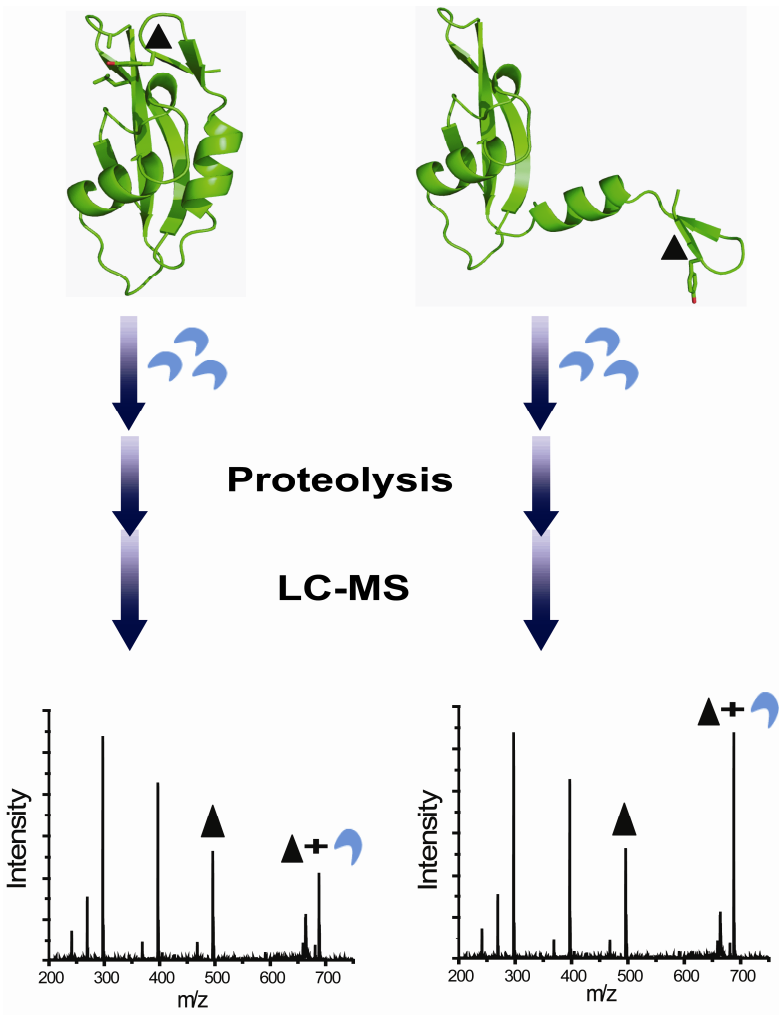


Figure 1.6 Analytical scheme showing how the abundances of the unmodified and modified peptides can identify amino acids that undergo changes in reactivity, which can then help determine regions in a protein that undergo a structural change.

Because of these potential problems, other approaches have been developed to measure amino acid modification extents. One method is to use isotopic labeling to determine relative reactivities of amino acids and has been used with lysine acetylation [37,38]. The relative reactivity of a given lysine residue can be determined by measuring the ion abundance ratio between the unlabeled modified peptide and the isotopically

labeled version of the modified peptide. Because an isotopically labeled version of the reagent is used, concerns over changes in ionization efficiency upon modification are minimized. Also, the isotopic labeling can be helpful for identifying modified peptides in a complex mixture.

A second approach to obtain more reliable information about the extent of modification is to use an internal standard, which is usually another peptide that does not contain a modifiable residue [39]. This method works well when comparing the differential reactivity of two or more residues of a given amino acid in the same protein. Typically, the ion abundance of the unmodified version of the peptide containing the modifiable residue of interest is compared to the ion abundances of peptides that do not contain modifiable residues. By measuring the ion abundance of the unmodified peptide, ionization efficiency concerns are minimized, and semi-quantitative measurements of the modification extent can be made.

A third approach developed to more reliably measure modification extents is to measure the modification reaction kinetics of the protein or, specifically, its peptide fragments that contain the modified residues of interest [31,40-42]. Rate coefficients can be determined by monitoring the extent of modification over a range of reaction times or reagent concentrations (Figure 1.7). These kinetic measurements can be performed by either measuring the increase in the ion abundance of the modified version of a peptide over time or the decrease in the ion abundance of the unmodified version of a peptide over time. Measuring modification kinetics has the additional potential advantage of providing more detailed structural information as most studies demonstrate that a good correlation exists between reaction rate coefficients and a residue's solvent accessibility.

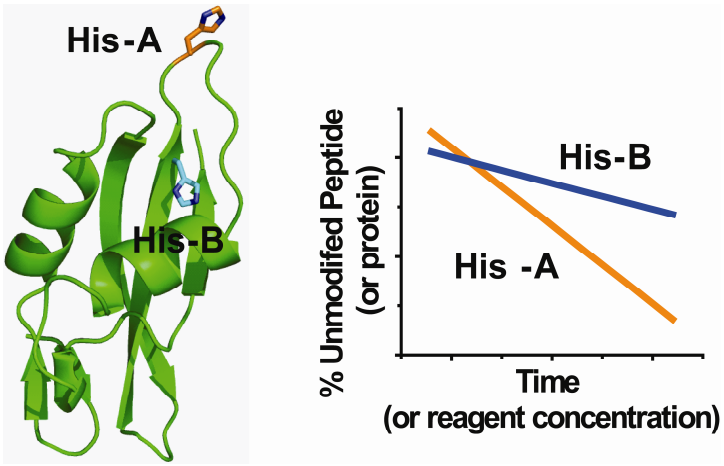


Figure 1.7 Plot of unmodified peptide (or protein) as a function of time (or reagent concentration) showing that a more buried residue His-B (blue) reacts more slowly and therefore has a shallower slope in the kinetics plot compared to a more exposed residue like His-A (orange).

1.5 Amino Acid Specific Labels and Examples of Their Applications

Amino acid specific reactions for only 8 of the 20 naturally occurring amino acids have been used to monitor protein structure in conjunction with MS. In general, the reactivity of the different functional groups in a protein relies on the nucleophilicity of the amino acid and the accessibility of the residue to the reagent. In the sections that follow, we describe amino acid specific reactions for Arg, His, and Lys with a particular focus on (i) the reaction mechanisms; (ii) the range of reaction conditions that are accessible with these reactions; (iii) any side reactions that need to be contended with; and (iv) specific biochemical problems that have been addressed. These three amino acids were chosen for the β2m structural studies described in the following chapters because of the predicted importance of these residues in Cu(II) binding and oligomer assembly, the suitability of the labels for studying proteins under native conditions, and the availability of relatively small labels for these particular residues.

1.5.1 Arginine

1.5.1.1 Reaction Mechanism

The modification of arginine residues is typically based on the reaction of vicinal dicarbonyl compounds to form cyclic adducts. For the structural studies of β2m, 2,3-butanedione (BD) was chosen because of its relatively small size. The reaction is shown in Figure 1.8.

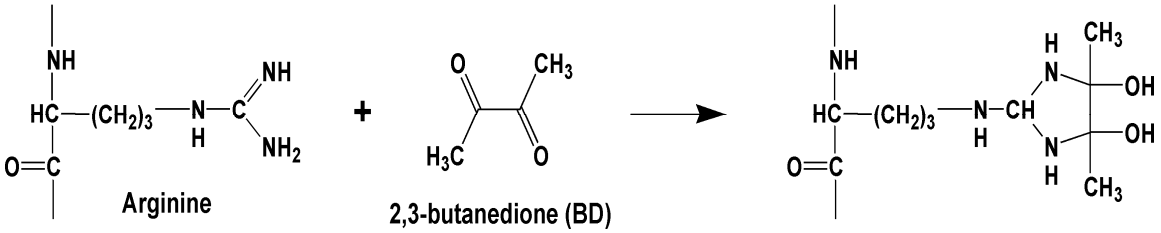


Figure 1.8 Modification reaction of arginine by 2,3-butandeione (BD).

1.5.1.2 Reaction Conditions

Previous studies show that arginine specific reactions are most commonly carried out between room temperature and 37 ^{o}C over a pH range of 7-10 with 50-100 mM buffer. The dicarbonyl labeling reagents are fairly soluble in water at concentrations up to 200 mM. Reagent concentrations typically vary from 0.1 to 3000-fold molar excess relative to the protein of interest, and reaction time of ~ 30 min. For BD, reactions are typically, but not universally, carried out in the dark to avoid possible photoactivation of this molecule, which could enhance nonspecific reactions with groups other than arginine [43,44]. Considered as a whole the accessible reactions conditions are almost ideal for studying proteins under native conditions. It has been noticed, however, that the adducts

that are formed with the various dicarbonyls can be somewhat unstable, especially at pHs below 7. Borate has been shown to stabilize the adducts and prevent the regeneration of arginine as shown in Figure 1.9 [45].

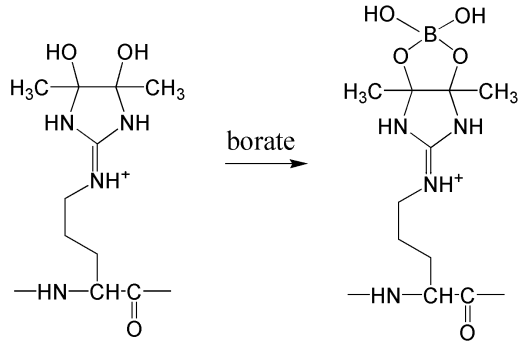


Figure 1.9 Reaction showing how borate can stabilize adducts that are formed by arginine with the various dicarbonyl compounds.

1.5.1.3 Biochemical Problems That Have Been Addressed

Arginine specific labeling combined with MS has been actively used to probe protein structures. Arginine has the highest pK_a of all amino acid side chains and thus is essentially always protonated under physiological conditions. As such arginines are almost always found on protein surfaces, so their value as targets for protein surface mapping would seem to be low because one would always expect to find them on a protein's surface. However, their ability to form intramolecular salt bridges with carboxylates, which could reduce their reactivity, can occasionally make them useful residues to probe. The use of 1,2 cyclohexanedione to modify lysozyme followed by mass spectrometric analysis was first reported by Przybylski and coworkers [40]. Only four of the eleven arginine residues in hen egg white lysozyme were modified, and the reactivities that were observed correlated inversely with the solvent accessibilities of

these arginine residues. The selective modification of only four of the arginines was ascribed to the presence of neighboring proton acceptor groups that lead to intramolecular catalysis, indicating that this reaction is very sensitive to the residue's local chemical environment.

Arginine modification and MS have also been employed to probe the active sites of proteins. For example, modification by BD with MALDI-TOF MS detection was used by Decottignies and co-workers to investigate the reactivities of the active site arginine residues of sorghum NADP-malate dehydrogenase [46]. MALDI-TOF mass spectrometry identified Arg87 and Arg134 as the sites of modification, indicating that these two arginine residues are the most reactive and accessible arginines in the active site.

An alternative covalent labeling method that uses BD and phenylboronic acid to cyclize the initial arginine adduct of BD was demonstrated by Leitner and Lindner [31]. Phenylboronic acid acted to stabilize the initial arginine adducts and improved the yield and detection efficiency of the modified proteins and peptides. [47]. The reactivities of the Arg residues in cytochrome c, lysozyme, and ubiquitin were still found to be in good agreement with the solvent accessibilities of these residues as obtained from crystal structures, indicating that the phenylboronic acid cap did not adversely affect the structural information gained from the labeling reaction. In addition, reactions with myoglobin, apomyoglobin and ribonuclease A before and after disulfide reduction were consistent with the known structural changes undergone by these proteins.

When an electrostatic interaction is suspected to be a major contributor to a protein-ligand binding interaction, arginine selective modification reactions can be a valuable probe because of arginine's ability to form salt bridges. Arginine specific

reagents have been used to examine the surface topology of the quaternary structure-dependent heparin-binding region of bovine seminal plasma PDC-109 by comparing the reactivities of arginine residues of the free and the heparin-bound protein [48]. Peptide mapping using N-terminal sequencing and ESI-MS revealed that Arg57, Arg64 and Arg104 were protected from modification in the heparin-bound protein. These Arg residues are suspected to be a part of the cationic face in the PDC-109 oligomeric state that binds tightly to heparin.

1.5.2 Histidine

1.5.2.1 Reaction Mechanism

Diethylpyrocarbonate (DEPC) has almost exclusively been used to modify histidine residues. A single modification predominates at low DEPC concentrations, but higher concentrations of the reagent can lead to a second carbethoxylation of histidine. The reactions of DEPC with histidine and other residues are shown in Figure 1.10.

1.5.2.2 Reaction Conditions

The reactions with DEPC are mostly carried out at room temperature and 37 ^{o}C in the pH range of 5.5-7.5. DEPC is insoluble in water, but it does dissolve in water-miscible solvents such as ethanol and acetonitrile. Typically, reagent amounts between 0.5 to 1000 fold molar excess are used with reaction times ranging from 1 min to 2h. These conditions are suitable for studying proteins under native conditions; however, because DEPC is not water soluble, care must be taken to ensure that the amount of organic solvent added does not disrupt the protein structure.

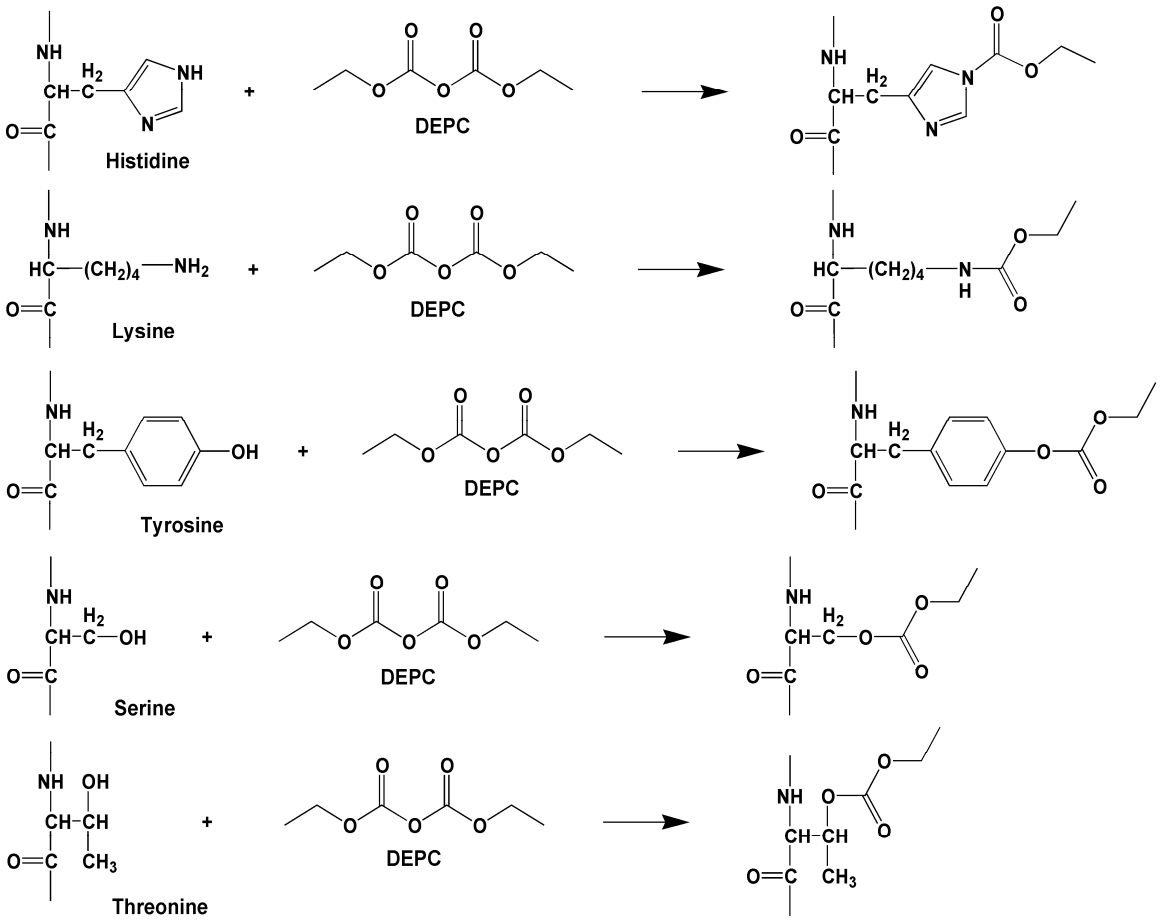


Figure 1.10 Modification reactions of diethylpyrocarbonate (DEPC) with histidine, lysine, tyrosine, serine, and threonine.

1.5.2.3 Side Reactions

Studies have shown that DEPC reacts with lysine at higher pHs. The reagent can also modify tyrosine, cysteine, arginine, serine and threonine residues with varying degrees of efficiencies [42,49].

1.5.2.4 Other Details of Reactivity

DEPC hydrolysis is known to occur with a reaction half-life of 9 min at 25 °C and pH 7, and the hydrolysis reaction is significantly faster at higher pHs [50]. Mono-modification of histidine residues is reversible under both acidic and alkaline conditions and in the presence of nucleophiles such as hydroxylamine and tris(hydroxymethyl)aminomethane (TRIS) [49]. The half-life of N-carbethoxyimidazole is about 55 hours at pH 7, 2 hours at pH 2, and 18 min a pH 10 [51]. Bis-modification of histidine residues, however, is irreversible. Modification of lysine and tyrosine residues are irreversible and reversible, respectively [49], and it has been recently shown that modifications to Ser and Thr residues are reversible over a 20 hr time period [42]. Finally, the extent of histidine's reactivity with DEPC is found to increase in the presence of acetate buffers in the pH range of 6 to 7 [52].

1.5.2.5 Biochemical Problems That Have Been Addressed

Histidine is the third least frequently occurring amino acid in proteins, yet its aromaticity, moderate basicity, H-bonding capacity, and ability to bind divalent transition metals causes it to be quite commonly involved in protein biochemistry. This fact and the ease with which DEPC can be used make histidine a relatively common target for studying protein surface structure. Most studies indicate that histidine's reactivity is controlled by its solvent accessibility and possibility its protonation state. Halsall and co-workers used MALDI-TOF and ESI-MS to determine the location and extent of DEPC modification of α_1-acid glycoprotein [53]. Results indicated that His97 was modified at a faster rate than His100, which was presumed to reflect differences in either the ionization

state of the imidazole or its accessibility to DEPC. This study also observed DEPC-modified lysine residues with the amount of lysine modification increasing over the pH range of 6 to 8. The reactivity of insulin with DEPC has also been investigated [52]. At DEPC:insulin ratios of 50:1, the products consisted of mixtures of the monomodified (CEt), formyl-CEt (FCEt) modified, and urethane-CEt (UCEt) modified histidine. Peptide mass mapping indicated that carbethoxylation occurred at His5, His10, and N-termini of both subunits. However, only biscarbethoxylation (FCEt and UCEt) of His10 was observed. The ε-nitrogen (NE) and the δ-nitrogen (ND) of His10 both have high solvent accessibility in both subunits, whereas His5 was only accessible on one side of the imidazolyl ring (Figure 1.11). These findings reveal that reactivity correlates well with the calculated solvent accessibilities of the imidazolyl nitrogen atoms.

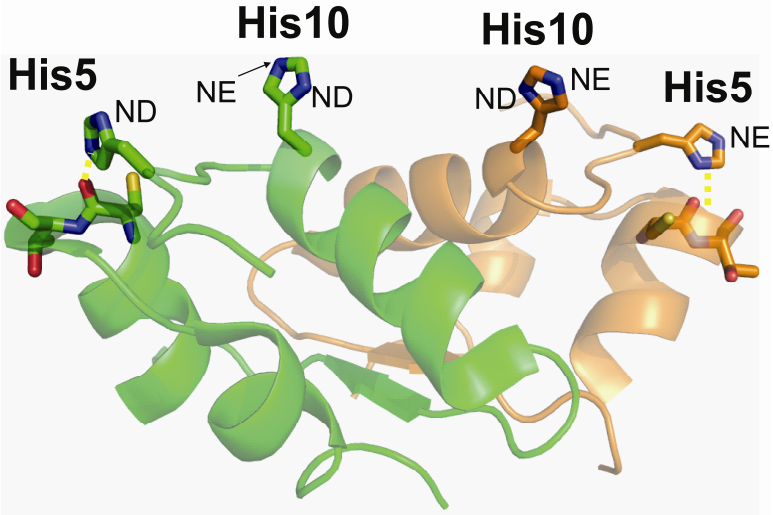


Figure 1.11 Crystal structure of the insulin dimer (PDB code 4INS). The ε-nitrogen (NE) and δ-nitrogen (ND) of His-10 are both accessible. His-5 participates in the insulin core structure and shows an accessible surface for either the ε-nitrogen (NE) or the δ-nitrogen (ND), depending on the monomer.

Several studies have also used DEPC modification to characterize protein-ligand interactions. These studies have involved modifying histidine residues then assessing the protein's activity or monitoring the reactivity of a protein's histidine residues with DEPC in the presence and absence of a ligand. Glocker and co-workers studied the role of histidine residues in the ligand-receptor interaction of rhM-CSFβ [39]. Hamasaki and co-workers investigated the possible roles of histidine residues in the conformational changes that occur during human erythrocyte band 3-mediated anion exchange [54].

An important role of histidine residues in proteins is to bind transition metals, and thus DEPC modification has been used to map metal-binding sites in proteins. A particularly noteworthy example involves an assessment of the Cu(II) binding sites of the prion protein (PrP) [55]. MALDI-MS data showed that five histidine residues within HuPrP23-98 were protected by Cu(II) from DEPC modification. However, in HuPrP23-98 solutions incubated with Zn(II) and Ni(II), only one or two histidines were partially protected from DEPC modification. These results validate the copper-binding specificity of the prion protein.

1.5.3 Lysine

1.5.3.1 Reaction Mechanism

Amino groups such as the ε-NH_2 of lysine residues and α-NH_2 of N-termini are most commonly modified by acylation with various organic acid anhydrides such as acetic anhydride, maleic anhydride, and succinic anhydride as well as *N*-hydroxysuccinimide derivatives. For the structural studies of β2m, sulfo-*N*-

hydroxysuccinimide acetate (NHSA) was chosen because of its relatively small size, solubility in water, and suitability for studying proteins under native conditions. The reaction of NHSA with lysine is shown in Figure 1.12.

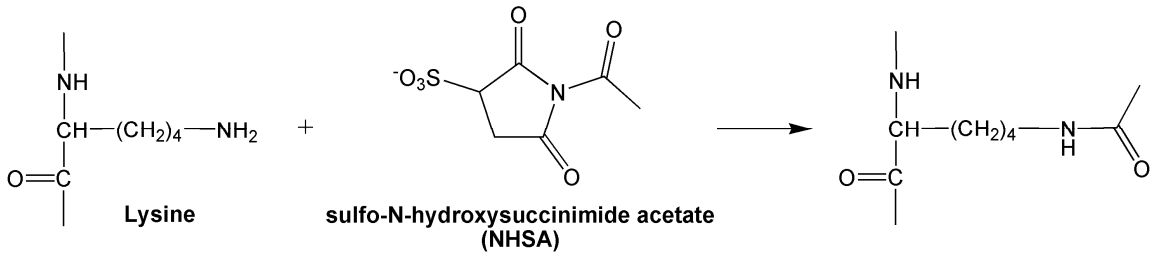


Figure 1.12 Modification reactions of lysine residues by sulfo-N-hydroxysuccinimide acetate (NHSA).

1.5.3.2 Reaction Conditions

The reactions of *N*-hydroxysuccinimide derivatives are typically carried out between 20 and 37 ^{o}C at pHs ranging from 5-8. NHSA is sparingly soluble in water. Reagent amounts vary from 5 to 5000 fold molar excess, and reaction times range from 2 to 80 min. An interesting aspect of NHSA is its ability to efficiently modify lysine residues at lower pH, which allows these covalent labels to be used under acid-unfolded conditions.

1.5.3.3 Biochemical Problems That Have Been Addressed

Lysine is one of the most common amino acids found in proteins, and it is almost always found on protein surfaces. Even so, modification of lysine residues has been the most commonly used covalent labeling strategy to probe protein surface structure. In one

of the earliest examples of using covalent labeling and MS, Knock et al. determined the relative reactivities of amine groups in Aplysia egg-laying hormone using the *N*-hydroxysuccinimide ester of biotin [56]. Sequence analysis and FAB-MS revealed that the α-amino group at the N-terminus was the least reactive.

Novak et al. used top-down Fourier transform ion cyclotron resonance mass spectrometry (FTICR-MS) combined with acetylation with N-hydroxysuccinimidyl acetate to study reactivities of the amino groups of ubiquitin under native and denaturing conditions [36]. Under denaturing conditions ubiquitin was fully modified with 21-fold molar excess of the label. Under native conditions, not all of the primary amino groups were modified even with an 81-fold molar excess. FTICR-MS data revealed the order of reactivities of the lysine residues, which are in good agreement with the solvent accessibility of residues obtained from the crystal structure of ubiquitin.

Zou and co-workers compared the reactivities of lysine residues of full-length native and hyperphosphorylated human replication protein A (RPA) using *N*-hydroxysuccinimide biotin [57]. MALDI-TOF MS and ESI-MS analysis revealed that Lys343 was protected from modification in the hyperphosphorylated RPA, indicating a structural rearrangement occurs upon phosphorylation. Alter and co-workers used the solvent accessibility of lysine residues to evaluate the relevance of three dimensional models of native RPA generated by remote-homology-based modeling procedures to the native solution-state structure of the protein [58]. The reaction rate constants of tryptic fragments with sulfo-N-hydroxysuccinimide were used to monitor the microenvironment of the lysine residues. These reactivities were then compared to the anticipated

reactivities of candidate structural models. Results showed that this approach is capable of assessing structural models and can be a basis for selecting the most relevant model.

Lysine residues are very commonly found on the surface of proteins, but they are not as commonly found at protein-protein interfaces, presumably due to the energetic cost of burying its positively-charged side chain. Nonetheless, the frequency of lysine residues on protein surfaces still make them useful probes of protein-protein interactions. Wang et al. characterized the interactions of rhodopsin and the α-subunit of transducin (G_t) using acetylation of lysine residues with sulfosuccinimidyl acetate [59]. Results showed that the modification extents of the lysine residues located in the cytoplasmic C1 and C2 loops and the C-terminal tail of rhodopsin were decreased upon light activation, indicating that these regions undergo conformational changes. In the presence of $G_t\alpha$(340-350), a synthetic peptide corresponding to the C-terminal residues of the α-subunit of G_t, acetylation sites on cytoplasmic loops 1, 2, and 4 of Rh* were protected, suggesting that these sites are involved in the interaction with G_t.

Heck and co-workers used a combination of lysine acetylation with N-acetyl-succinimide and nanoLC-MALDI MS to probe the surface interactions of the DNase domain of Colicin E9 (E9) with its immunity protein Im9 [60]. This system was used to test if chemical modification and nanoLC-MALDI MS together with data filtering using immonium marker ions could be used to characterize protein-protein interactions. In the presence of Im9, Lys55, Lys63, Lys76, Lys81, Lys89 and Lys97 of E9 were protected from modification (Figure 1.13). Based on the crystal structure of the E9:Im9 complex, Lys89 and Lys97 are known to be directly involved in the interaction site. Lys76 and Lys81 are located in a α-helical region near the interaction surface, and Lys55 and Lys63

do not interact with Im9. These data suggest that Im9 binding to E9 induces conformational changes in some amino acid residues. The results again illustrate that chemical modification approaches can be used to probe protein-protein interactions and provide complementary data to already available structural models from X-ray crystallography.

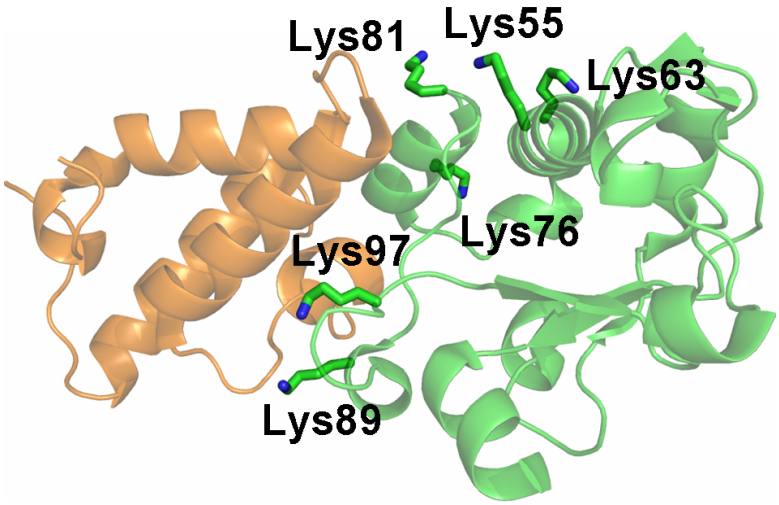


Figure 1.13 Crystal structure of the E9:Im9 complex (PDB code 1EMV). Lys89 and Lys97 are directly involved in the interaction. Lys76 and Lys81 are located in a α-helical region near the interaction surface. Lys55 and Lys63 do not interact with Im9 (orange), and the region of E9 (green) containing these residues possibly undergoes a conformational change upon Im9 binding.

1.6 Summary

The use of selective covalent labeling in conjunction with MS has been extensively used in recent years to study protein structure and interactions. As shown in the previous section, the information obtained from this methodology is complementary to other biophysical techniques. This dissertation presents improvements in this approach

and its application in determining the metal binding site, binding-induced conformational changes, and protein-protein interfaces of the oligomers that precede amyloid formation of β2m.

Chapter 2 describes an improved protein surface mapping using DEPC with mass spectrometric detection. It also details the use of measuring the reaction kinetics for ensuring protein structural integrity during covalent labeling reactions.

In Chapter 3, our improved mass spectrometry-based protein surface mapping method is used to determine the Cu(II) binding site and any Cu(II)-induced conformational changes of β2m.

Chapter 4 shows how covalent labeling in conjunction with mass spectrometry is used to identify the amino acid residues that mediate β2m dimer formation.

Chapter 5 describes the use of the protein surface mapping approach to gain insight into the structure of the pre-amyloid tetramer of β2m.

In Chapter 6, the differences in oligomerization and/or fibril formation of β2m in the presence of different divalent metal ions is investigated.

Finally, Chapter 7 contains conclusions and future research directions.

1.7 References

1. Dobson, C.M. (1999) Protein misfolding, evolution and disease. *Trends Biochem. Sci. 24*, 329-330.

2. Kelly, J.W. (1998) The alternative conformations of amyloidogenic proteins and their multistep assembly pathways. *Curr. Opin. Biol. 8*, 101-106.

3. Zerovnik, E. (2002) Amyloid-fibril formation: proposed mechanisms and relevance to conformational disease. *Eur. J. Biochem. 269*, 3361-3371.

4. Stefani, M. (2004) Protein misfolding and aggregation: new examples in medicine and biology of the dark side of the protein world. *Biochim. Biophys. Acta 1739*, 5-25.

5. Sunde, M., Serpell, L.C., Bartlam, M., Fraser, P.E., Pepys, M.B., and Blake, C.C. (1997) Common core structure of amyloid fibrils by synchroton X-ray diffraction. *J. Mol. Biol. 273*, 729-739.

6. López de la Paz, M., Goldie, K., Zurdo, J., Lacroix, E., Dobson, C.M., Hoenger, A., and Serrano, L. (2002) *De novo* designed peptide-based amyloid fibrils. *Proc. Natl. Acad. Sci. USA 99*, 16052-16057.

7. Tycko, R. (2004) Progress towards a molecular-level structural understanding of amyloid fibrils. *Curr. Opin. Struct. Biol. 14*, 96-103.

8. Rochet, J.C., and Lansbury, P.T. Jr. (2000) Amyloid fibrillogenesis: themes and variations. *Curr. Opin. Struct. Biol. 10*, 60-68.

9. Floege, J., and Ehlerding, G. (1996) Beta-2-microglobulin associated amyloidosis. *Nephron 72*, 9-26.

10. Bjorkman, P.J., Saper, M.A., Samraoui, B., Bennett, W.S., Strominger, J.L., and Wiley, D.C. (1987) Structure of the human class-I histocompatibility antigen, HLA-A2. *Nature 329*, 506-512.

11. Keating, M.J. (1999) Chronic lymphocytic leukemia. *Semin. Oncol. 26*, 107-114.

12. Malaguarnera, M., Restuccia, S., Di Fazio, I., Zoccolo, A.M., Trovato, B.A., and Pistone, G. (1997) Serum beta-2-microglobulin in chronic hepatitis C. *Dig. Dis. Sci. 42*, 762-766.

13. Okon, M., Bray, P., and Vucelic, D. (1992) H^1-NMR assignments and secondary structure of human beta-2-microglobulin in solution. *Biochemistry 31*, 8906−8915.

14. Morgan, C.J., Gelfans, M., Atreya, C., and Miranker, A.D. (2001) Kidney dialysis-associated amyloidosis: a molecular role for copper in fiber formation. *J. Mol. Biol. 309*, 339-345.

15. Villanueva, J., Hoshino, M., Katou, H., Kardos, J., Hasegawa, K., Naiki, H., and Goto, Y. (2004) Increase in the conformational flexibility of β_2-microglobulin upon copper binding: a possible role for copper in dialysis-related amyloidosis. *Protein Sci. 13*, 797-809.

16. Calabrese, M.F., and Miranker, A.D. (2007) Formation of a stable oligomer of β-2 microglobulin requires only a transient encounter with Cu(II). *J. Mol. Biol. 367*, 1-7.

17. Eakin, C.M., Berman, A.J., and Miranker, A.D. (2006) A native to amyloidogenic transition regulated by a backbone trigger. *Nature Struct. Mol. Biol. 13*, 202-208.

18. Antwi, K., Mahar, M., Tyson, J.F., and Vachet, R.W. (2008) Cu(II) organizes β-2-microglobulin oligomers but is released before amyloid formation. *Protein Sci. 17*, 748-759.

19. Eakin, C.M., Attenello, F.J., Morgan, C.J., and Miranker, A.D. (2004) Oligomeric assembly of native-like precursors precedes amyloid formation by β-2-microglobulin. *Biochemistry 43*, 7808-7815.

20. Walsh, D.M., Lomakin, A., Benedek, G.B., Condron, M.M., and Teplow, D.B. (1997) Amyloid beta-protein fibrillogenesis – detection of a protofibrillar intermediate. *J. Biol. Chem. 272*, 22364-22372.

21. Friedhoff, P., von Bergen, M., Mandelkow, E.M., Davies, P., and Mandelkow, E. (1998) A nucleated assembly mechanism of Alzheimer paired helical filaments. *Proc. Natl. Acad. Sci. USA 95*, 15712-15717.

22. Ferrao-Gonzales, A.D., Robbs, B.K., Moreau, V.H., Ferreira, A., Juliano, L., Valente, A.P., Almeida, F.C.L., Silva, J.L., and Foguel, D. (2005) Controlling beta-amyloid oligomerization by the use of naphthalene sulfonates – trapping low molecular weight oligomeric species. *J. Biol. Chem. 280*, 34747-34754.

23. Smith, A.M., Jahn, T.R., Ashcroft, A.E., and Radford, S.E. (2006) Direct observation of oligomeric species formed in the early stages of amyloid fibril formation using electrospray ionization mass spectrometry. *J. Mol. Biol. 364*, 9-19.

24. Bucciantini, M., Giannoni, E., Fabrizio, C., Baroni, F., Formigli, L., Zurdo, J., Taddei, N., Ramponi, G., Dobson, C.M., and Stefani, M. (2002) Inherent toxicity of aggregates implies a common mechanism for protein misfolding diseases. *Nature 416*, 507-511.

25. Mendes Sousa, M., Cardoso, I., Fernandes, R., Guimaraes, A., and Saraiva, M.J. (2001) Deposition of transthyretin in early stages of familial amyloidotic polyneuropathy. Evidence for toxicity of nonfibrillar aggregates. *Am. J. Pathol. 159*, 1993-2000.

26. McParland, V.J., Kad, N.M., Kalverda, A.P., Brown, A., Kirwin-Jones, P., Hunter, M.G., Sunde, M., and Radford, S.E. (2000) Partially unfolded states of β_2-microglobulin and amyloid formation in vitro. *Biochemistry 39*, 8735-8746.

27. Verdone, G., Corazza, A., Viglino, P., Pettirossi, F., Giorgetti, S., Mangione, P., Andreola, A., Stoppini, M., Bellotti, V., and Esposito, G. (2002) The solution structure of human β2-microglobulin reveals the prodromes of its amyloid transition. *Protein Sci. 11*, 487-499.

28. Trinh, C.H., Smith, D.P., Kalverda, A.P., Phillips, S.E.V., and Radford, S.E. (2002) Crystal structure of monomeric human beta-2-microglobulin reveals clues to its amyloidogenic properties. *Proc. Natl. Acad. Sci.USA 99*, 9771-9776.

29. Yamaguchi, K., Katou, H., Hoshino, M., Hasegawa, K., Naiki, H., and Goto, Y. (2004) Core and heterogeneity of β_2-microglobulin amyloid fibrils as revealed by H/D exchange. *J. Mol. Biol. 338*, 559-571.

30. Kad, N.M., Myers, S.L., Smith, D.P., Smith, D.A., Radford, S.E., and Thomson, N.H. (2003) Hierarchical assembly of β_2-microglobulin amyloid in vitro revealed by atomic force microscopy. *J. Mol. Biol. 330*, 785-797.

31. Leitner, A., and Lindner, W. (2005) Functional probing of arginine residues in proteins using mass spectrometry and an arginine-specific covalent tagging concept. *Anal. Chem. 77*, 4481-4488.

32. Kim, J.Y., Kim, K.W., Kwon, H.J., Lee, D.W., and Yoo, J.S. (2002) Probing lysine acetylation with a modification-specific marker ion using high-performance liquid chromatography/electrospray-mass spectrometry with collision-induced dissociation. *Anal. Chem. 74*, 5443-5449

33. Gabant, G., Augier, J., and Armengaud, J. (2008) Assessment of solvent residues accessibility using three Sulfo-NHS-biotin reagents in parallel: application to footprint changes of a methyltransferase upon binding its substrate. *J. Mass Spectrom. 43*, 360–370.

34. Meng, F.Y., Forbes, A.J., Miller, L.M., and Kelleher, N.L. (2005) Detection and localization of protein modifications by high resolution tandem mass spectrometry. *Mass Spectrom. Rev. 24*, 126-134.

35. Zhai, H.L., Dorrestein, P.C., Chatterjee, A., Begley, T.P., and McLafferty, F.W. (2005) Simultaneous kinetic characterization of multiple protein forms by top down mass spectrometry. *J. Am. Soc. Mass Spectrom. 16*, 1052-1059.

36. Novak, P., Kruppa, G.H., Young, M.M., and Schoeniger, J. (2004) A top-down method for the determination of residue-specific solvent accessibility in proteins. *J. Mass Spectrom. 39*, 322-328.

37. Hochleitner, E.O., Borchers, C., Parker, C., Bienstock, R.J., and Tomer, K.B. (2000) Characterization of a discontinuous epitope of the human immunodeficiency virus (HIV) core protein p24 by epitope excision and differential chemical modification followed by mass spectrometric peptide mapping analysis. *Protein Sci. 9*, 487-496.

38. Glocker, M.O., Borchers, C., Fiedler, W., Suckau, D., and Przybylski, M. (1994) Molecular characterization of surface topology in protein tertiary structures by amino-acylation and mass spectrometric peptide mapping. *Bioconjugate Chem. 5*, 583-590.

39. Glocker, M.O., Kalkum, M., Yamamoto, R., and Schreurs, J. (1996) Selective biochemical modification of functional residues in recombinant human macrophage colony-stimulating factor β (rhM-CSF β): Identification by mass spectrometry. *Biochemistry 35*, 14625-14633.

40. Suckau, D., Mak, M., and Przybylski, M. (1992) Protein surface topology-probing by selective chemical modification and mass spectrometric peptide mapping. *Proc. Natl. Acad. Sci. USA 89*, 5630-5634.

41. Gao, Y., and Wang, Y. (2006) Site-selective modifications of arginine residues in human hemoglobin induced by methylglyoxal. *Biochemistry 45*, 15654-15660.

42. Mendoza, V.L., and Vachet, R.W. (2008) Protein surface mapping using diethylpyrocarbonate with mass spectrometric detection. *Anal. Chem. 80*, 2895-2904.

43. Fliss, H., and Viswanatha, T. (1979) 2,3-Butanedione as a photosensitizing agent – Application to alpha-amino acids and alpha-chymotrypsin. *Can. J. Biochem. 57*, 1267-1272.

44. Riordan, J.F. (1973) Functional arginyl residues in carboxypeptidase-A – Modification with butanedione. *Biochemistry 12*, 3915-3923.

45. Riordan, J.F. (1979) Arginyl residues and anion binding-sites in proteins. *Mol. Cell Biochem. 26*, 71–92.

46. Schepens, I., Ruelland, E., Miginiac-Maslow, M., Marechal, P., and Decottignies, P. (2000) The role of active site arginines of Sorghum NADP-malate dehydrogenase in thioredoxin-dependent activation and activity. *J. Biol. Chem. 275*, 35792-35798.

47. Leitner, A., Amon, S., Rizzi, A., and Lindner, W. (2007) Use of the arginine-specific butanedione/phenylboronic acid tag for analysis of peptides and protein digests using matrix-assisted laser desorption/ionization mass spectrometry. *Rapid Comm. Mass Spectrom. 21*, 1321-1330.

48. Calvete, J.J., Campanero-Rhodes, M.A., Raida, M., and Sanz, L. (1999) Characterization of the conformational and quaternary structure-dependent heparin-binding region of bovine seminal plasma protein PDC-109. *FEBS Letters 444*, 260-264.

49. Miles, E.W. (1977) Modification of histidyl residues in proteins by diethylpyrocarbonate. *Method Enzymol. 47*, 431-442.

50. Lundblad, R.L., and Noyes, C.M. (1984) Chemical reagents for protein modification. CRC Press, Boca Raton Fl. Ch. 1, 105-126.

51. Melchoir, W.B. Jr, and Fahrney, D. (1970) Ethoxyformylation of proteins. Reaction of ethoxyformic anhydride with α-chymotrypsin, pepsin, and pancreatic ribonuclease at pH 4. *Biochemistry 9*, 251-258.

52. Kalkum, M., Przybylski, M., and Glocker, M.O. (1998) Structure characterization of functional histidine residues and carbethoxylated derivatives in peptides and proteins by mass spectrometry. *Bioconjugate Chem. 9*, 226-235.

53. Dage, J.L., Sun, H., and Halsall, H.B. (1998) Determination of diethylpyrocarbonate-modified amino acid residues in α_1-acid glycoprotein by high-performance liquid chromatography electrospray ionization-mass spectrometry and matrix-assisted laser desorption/ionization-time of flight mass spectrometry. *Anal. Biochem. 257*, 176-185.

54. Jin, X.R., Abe, Y., Li, C.Y., and Hamasaki, N. (2003) Histidine-834 of human erythrocyte band 3 has an essential role in the conformational changes that occur during the band 3-mediated anion exchange. *Biochemistry 42*, 12927-12932.

55. Qin, K., Yang, Y., Mastrangelo, P., and Westaway, D. (2002) Mapping Cu(II) binding sites in prion protein by diethyl pyrocarbonate modification and matrix-assisted laser desorption ionization-Time of Flight (MALDI-TOF) mass spectrometric footprinting. *J. Biol. Chem. 277*, 1981-1990.

56. Knock, S.L., Miller, B.T., Blankenship, J.E., Nagle, G.T., Smith, J.S., and Kurosky, A. (1991) N-acylation of aplysia egg-laying hormone with biotin. *J. Biol. Chem. 266*, 24413-24419.

57. Liu, Y.Y., Kvaratskhelia, M., Hess, S., Qu, Y.X., and Zou, Y. (2005) Modulation of replication protein A function by its hyperphosphorylation-induced conformational change involving DNA binding domain B. *J. Biol. Chem. 280*, 32775-32783.

58. Nuss, J.E., Sweeney, D.J., and Alter, G.M. (2006) Reactivity-based analysis of domain structures in native replication protein A. *Biochemistry 45*, 9804-9818.

59. Wang, X., Kim, S.H., Ablonczy, Z., Crouch, R.K., and Knapp, D.R. (2004) Probing rhodopsin-transducin interactions by surface modification and mass spectrometry. *Biochemistry 43*, 11153-11162.

60. Scholten, A., Visser, N.F.C., van den Heuvel, R.H.H., and Heck, A.J.R. (2006) Analysis of protein-protein interaction surfaces using a combination of efficient lysine acetylation and nanoLC-MALDI-MS/MS applied to the E9:Im9 bacteriotoxin-immunity protein complex. *J. Am. Soc. Mass Spectrom. 17*, 983-994.

CHAPTER 2

IMPROVED PROTEIN SURFACE MAPPING USING DIETHYLPYROCARBONATE WITH MASS SPECTROMETRIC DETECTION

This chapter is part of a paper published as: Mendoza, V. L., and Vachet, R. W. (2008) Protein surface mapping using diethylpyrocarbonate with mass spectrometric detection. *Anal. Chem. 80*, 2895-2904.

2.1 Introduction

Determination of the molecular structure of proteins continues to be important in molecular biology because of the relationship between protein structure and function. Detailed structural information about a protein's spatial arrangement is usually obtained by X-ray crystallography and NMR; however, both of these techniques have some limitations. As such, methods based on mass spectrometry (MS) that overcome some of the shortcomings associated with these techniques are being developed to study higher-order protein structure.

MS does not directly provide amino-acid level information about 3D protein structure in solution because mass spectral measurements occur in the gas phase. Hence, some means of encoding this structural information in the mass-to-charge (m/z) ratio of the measured ions must be applied. Two general approaches have been used, non-covalent and covalent labeling, to accomplish this. Hydrogen deuterium exchange (HDX) coupled with MS has been extensively and successfully used as a non-covalent labeling tool to study protein structure and dynamics in solution [1-3]. HDX provides information about a protein's backbone structure, but back-exchange during the various stages of analysis (e.g. LC, MS/MS) can result in the loss of pertinent data and can complicate data interpretation. In contrast, covalent labeling techniques reports on protein side chains,

which can be complementary to HDX methods, while avoiding issues of back exchange. One kind of covalent labeling approach relies on reagents that irreversibly modify either specific amino acids [4-6] or most amino acids (e.g. hydroxyl radicals [7,8]). These reagents provide information about protein structure by identifying solvent exposed amino acids. The residue specific labels have the advantage of simplicity, requiring reagents and conditions that are readily accessible and easy to use, while methods based on hydroxyl radicals require special radiation sources to produce the reactive species.

While covalent labels usually provide information that is complementary to HDX methods, they have some advantages that make covalent labeling methods worth further investigation. First, the possibilities of back-exchange and scrambling are basically non-existent. Second, confident detection and identification of modified residues and/or regions of a protein can be readily done with most types of mass spectrometers because of the greater mass shifts that occur upon labeling. Finally, observation of small and locally restricted structural changes in proteins is potentially more straightforward because the labels can typically be localized to single amino acids. The main drawback of using covalent labels in comparison to HDX is that protein structure is more likely to be perturbed by these relatively large labels than when D is used to probe structure.

Given its simplicity and advantages, amino acid-specific covalent labeling of proteins with detection by MS has been used to map protein surfaces, identify ligand-binding sites, study protein-protein and protein-nucleic acid complexes, and detect ligand-induced conformational changes [9-18]. Because protein conformational changes affect their surface topology, amino acid residues involved in the structural changes can be identified from differential modification patterns. Information from surface mapping

experiments is reliable, though, only if the structural integrity of a protein is preserved during the reaction. This is a serious concern because of the relatively large size of typical amino acid-specific labels. Hence, appropriate checks are required to ensure that the labeling reaction does not distort the protein's structure and thus provide incorrect information. Despite the importance of such checks, a survey of over 60 publications that use MS and amino acid-specific labeling indicates that over 60% of these studies did nothing to ensure the structural integrity of the studied protein. About 33% of these studies either used circular dichroism (CD) spectroscopy or activity assays to check protein structure. However, these approaches are probably not sufficient to identify small-scale or local structural changes. CD spectroscopy provides only the population-weight average properties of a protein sample and may not be sensitive to local regions of a protein, and activity assays will only change significantly if a protein's structure around its active site is sufficiently perturbed. About 5% of the reports used fluorescence spectroscopy, which can provide information about local protein structural changes but only around tryptophan residues. Finally, only one of the 60+ articles limited the number of modifications to 1 per protein molecule, but this approach makes detection of the modified residues more difficult.

In this chapter, an improved amino acid-specific modification and MS-based approach for protein surface mapping is developed to address the need for a reliable approach that both ensures protein structural integrity during labeling experiments and provides readily detectable modifications. We demonstrate that measuring the kinetics of the covalent labeling reactions ensures that the covalent probe does not disrupt a protein's structure during the labeling reaction. The reaction rate coefficients determined also

provides a quantitative basis for more fully understanding the factors that influence amino acid reactivity. In addition, we show that diethylpyrocarbonate, which is a very convenient and easy-to-use label, may have some promise as a general covalent label because of its ability to probe up to 25% of the residues in the average protein.

2.2 Experimental Procedure

2.2.1 Materials

Human β-2-microglobulin (β2m) was obtained from Fitzgerald Industries International, Inc. (Concord, MA). Diethylpyrocarbonate (DEPC), imidazole, dithiothreitol (DTT), equine heart cytochrome c, and equine skeletal muscle myoglobin were obtained from Sigma-Aldrich (St. Louis, MO). Tris(hydroxymethyl)-aminomethane (Tris) and tris(hydroxymethyl)aminomethane hydrochloride (Tris-HCl) were purchased from EM Science (Gladstone, NJ). Ammonium acetate, methanol, acetonitrile, and acetic acid were obtained from Fisher Scientific (Fair Lawn, NJ). Trypsin was from Promega (Madison, WI). Centricon molecular weight cutoff (MWCO) filters were obtained from Millipore (Burlington, MA). Deionized water was prepared from a Millipore (Burlington, MA) Simplicity 185 water purification system.

2.2.2 DEPC Modification

The covalent modification reactions were performed for 1 min at 37 ^{o}C with 100 μM protein and 50 mM ammonium acetate at pH 7 and were initiated by adding various molar excesses of DEPC in acetonitrile. The total reaction volume was 100 μL for cytochrome c and myoglobin, and the total amount of acetonitrile was 1 %. The DEPC

reactions of β2m were performed for 1 min at 37 ^{o}C with 100 μM protein, 200 mM potassium acetate and 25 mM MOPS at pH 7.4 and were initiated by adding various molar excesses of DEPC. The total reaction volume for the experiments with β2m was 50 μL, and the total amount of acetonitrile was 1%. All the covalent modification reactions were quenched by adding 10 mM imidazole. The DEPC-treated samples were then purified using a 10,000 MWCO filter and reconstituted with deionized water to a final concentration of 250 μM.

2.2.3 Proteolytic Digestion

Before the addition of trypsin, 80 μL solutions containing 1 μg/μL of unmodified or DEPC-modified proteins in 50 mM Tris-HCl (pH = 7) and 5 mM $CaCl_2$ were incubated with 10 μL of acetonitrile at 40 ^{o}C for 45 minutes. Also, to reduce the disulfide bonds in myoglobin and β2m, the proteins were reacted with 10 mM DTT at 40 ^{o}C for 45 minutes prior to addition of the acetonitrile. Trypsin (0.5 μg/μL) was then added to yield a final enzyme:substrate ratio of 1:20. For the β2m samples, chymotrypsin (0.5 μg/μL) was added together with the trypsin. All samples were incubated at 37 ^{o}C for 16 hours. The enzymes were inactivated by adding 2 μL of acetic acid, and the samples were immediately frozen at -10 ^{o}C and analyzed within 24 hours.

2.2.4 HPLC Analysis

An HP1100 (Agilent, Wilmington, DE) HPLC system with a C18 column (15 cm x 2.1 mm, 5 μm particle size, Supelco, St. Louis, MO) was used for all LC experiments. Tryptic fragments were eluted using a linear gradient of methanol containing 0.1% acetic

acid that increased from 5% to 90% methanol over 30 min at a flow rate of 0.250 mL/min. The LC effluent was split at a ratio of 1:4 with the smaller fraction of the split flow being fed into the mass spectrometer.

2.2.5 Mass Spectrometry

Mass spectra were acquired on a Bruker Esquire-LC (Billerica, MA) quadrupole ion trap mass spectrometer equipped with an electrospray ionization source. Typically, the electrospray needle voltage was kept at 3-3.5 kV, and the capillary temperature was set to 300 ^{o}C. A voltage of 30-40 V was applied to skimmer 1, and the capillary offset voltage was between 50 and 60 V. For direct injection experiments similar source conditions were used, but sample was delivered at 1 μL/min using a syringe pump. Tandem mass spectra were acquired using isolation widths of 1.0 Da and excitation voltages between 0.6 and 1.0 V. Peptide sequences were determined from the MS/MS data via *de novo* sequencing or with the help of BioTools^{TM} (Bruker Daltonics, Billerica, MA). The sequence coverage for cytochrome c and β2m was 100%, while the sequence coverage for myoglobin was approximately 90%.

2.2.6 Determination of Solvent Accessibility

The 3D structures of cytochrome c and myoglobin were examined using the structures from the Protein Data Bank (PDB) (1AKK for cytochrome c and 1DWR for myoglobin). Solvent accessibility was calculated using GETAREA (http://curie.utmb.edu/getarea.html) [19]. A probe radius of 1.4 Å, representing the van der Waals sphere of water, was used.

2.2.7 Fluorescence Measurements

Fluorescence spectra were acquired with a Photon Technology International Quantamaster-4 SE. Tryptophan fluorescence measurements were obtained using slit widths of 2 nm and an excitation wavelength of 280 nm. Protein solutions (100 μM, 50 mM ammonium acetate at pH 7) were initially incubated at 37 ^{o}C for 30 min prior to addition of DEPC. Emission scans from 300-390 nm were taken at different time points after the addition of DEPC. Fluorescence intensity measurements immediately before and 1 to 5 min after the addition of DEPC were performed to determined the effect of DEPC modification on protein structure. The fluorescence measurements are reported as Average Emission Fluorescence Wavelength, $\langle\lambda\rangle$, as shown in equation 1 [20]:

$$\langle\lambda\rangle = \frac{\sum_{i=1}^{N} (I_i \lambda_i)}{\sum_{i=1}^{N} (I_i)} \qquad \text{Equation (1)}$$

2.2.8 Circular Dichroism (CD) Experiments

CD experiments were performed at 37 ^{o}C using a Jasco-710 spectropolarimeter. All experiments were done in 50 mM ammonium acetate at pH 7. Spectra were measured with a sample cell having a 0.1 cm path length and at a scan resolution of 0.2 nm, a scan rate of 50 nm/min, and a response time of 4 s.

2.3 Results and Discussion

2.3.1 Carbethoxylation of Histidine, Tyrosine, Serine and Threonine

Diethylpyrocarbonate (DEPC) reacts somewhat specifically with histidine residues at near-physiological pHs (Figure 2.1), but it can also react with tyrosine, cysteine, and lysine residues [9,21,22]. Covalent modification experiments with DEPC are usually performed for 30-120 min at pH ~ 7 [21-23]. In our experiments a 1 min reaction time was chosen to minimize the effect of DEPC hydrolysis on the kinetics of the modification reactions. DEPC hydrolysis is known to occur with a reaction half-life of 9 min at 25 °C and pH = 7 [24]. In addition, a short reaction time was used to improve the temporal resolution of the modification procedure and decrease the possibility of protein structural changes during modification. The reagent dose was also kept relatively low to minimize any structural changes caused by the reagent.

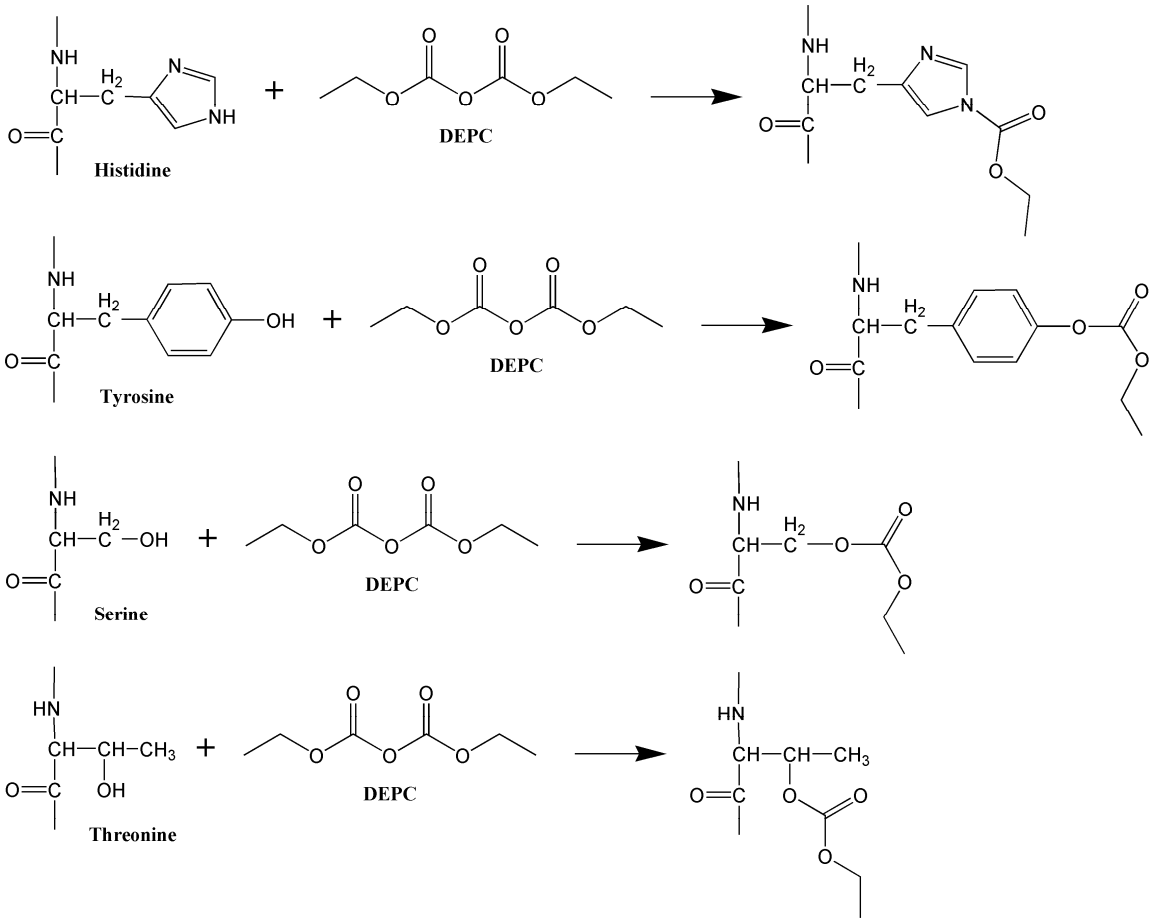


Figure 2.1 Reactions of histidine, tyrosine, threonine, and serine with DEPC.

Under our reaction conditions, LC-MS/MS data (e.g. Figure 2.2) show that many solvent exposed His, Tyr, Ser and Thr residues are modified by DEPC (Tables 2.1 and 2.2). Most of the modified amino acids have solvent accessible surface area (SASA) percentages above 30. Figure 2.2 shows the MS/MS spectra of the 16 modified peptides which confirm the modification sites that are reported in Tables 2.1 and 2.2. These spectra also demonstrate that modifications can be readily determined with single amino acid resolution. For example, unmodified y_3, y_4, y_5, and y_6 product ions along with modified y_7 ion indicate that His13 is the modified residue in the Ser11-Lys19 fragment

of β2m (Figure 2.2B). Similarly, unmodified b ions from b_9 through b_{14}, a modified b_{15} ion, and a complete series of modified y_2 to y_{13} ions indicate that Ser117 is the modified residue in the Tyr103-Lys119 fragment of myoglobin (Figure 2.2Q).

Table 2.1 DEPC modification rate coefficients for histidine residues.

Protein	Residue	% SASA Ratio[a]	SASA of Nε2 (Å^2)[a]	pK_a[b]	k[c] (M^{-1} s^{-1})
Cytochrome C	His18	15	3.9	4.3	0
	His26	34	6.8	6.5	0.069 ± 0.004
	His33	49	0.3	6.4	0
Myoglobin	His24	3	0	4.6	0
	His36	33	11.7	6.1	0.070 ± 0.007
	His48	65	6.4	6.7	0.067 ± 0.003
	His64	27	5.9	4.5	0.009 ± 0.001
	His81	89	18.8	7.4	0.053 ± 0.002
	His82	45	0	7.1	0
	His93	35	17.1	4.4	0
	His97	43	13.5	6.4	0.036 ± 0.003
	His113	43	9.9	6.2	0
	His116	44	0.6	6.4	0
	His119	20	0	6.2	0
β2m	His13	56	15.7	7.4	0.041 ± 0.003
	His31	25	11.5	7.2	0.010 ± 0.001
	His51	44	10.2	6.2	0.036 ± 0.003
	His84	0	0	7.3	0

[a] Solvent accessible surface area (SASA) calculated using GETAREA 1.1. 1.4 Å is used as the probe radius, and the calculated SASA is compared to the surface area of the side chain in a Gly-X-Gly tripeptide generating a % ratio. Residues with % ratio below 20 are typically considered buried whereas those with ratios above 30% are regarded as solvent exposed

[b] pK_a calculated using PROPKA

[c] Rate coefficients from dose-response plots like those shown in Figure 2.5

DEPC is known to react with Tyr residues in addition to His residues, but the modification of Ser and Thr residues is previously unreported. Like Tyr residues, Ser and Thr have weakly nucleophilic alcohol groups capable of reacting with DEPC. We find that the reactivity of Ser and Thr residues is mostly limited, though, to those residues that are highly exposed (i.e. SASA % ratios > 50%), and so only 21% and 5% of the Ser and Thr residues, respectively, are modified in the three proteins studied here. Nonetheless, this observation dramatically expands the potential of DEPC as a probe of protein structure. Histidine and tyrosine residues cover only 2.2% and 3.3%, respectively, of the sequence of a typical protein [25]. In contrast, Ser and Thr cover 7.4% and 6.0% of the sequence of a typical protein [25]. Indeed, Ser is the third most abundant residue in proteins. Because Lys (5.8% frequency) and Cys (1.8% frequency) can also be modified by DEPC, this reagent has the potential to cover about 25% of the sequence of a typical protein. This degree of coverage could make DEPC a very general surface mapping reagent for those proteins having typical numbers of His, Tyr, Ser, Thr, Lys, and Cys residues.

Table 2.2 DEPC modification rate coefficients for tyrosine, serine and threonine residues.

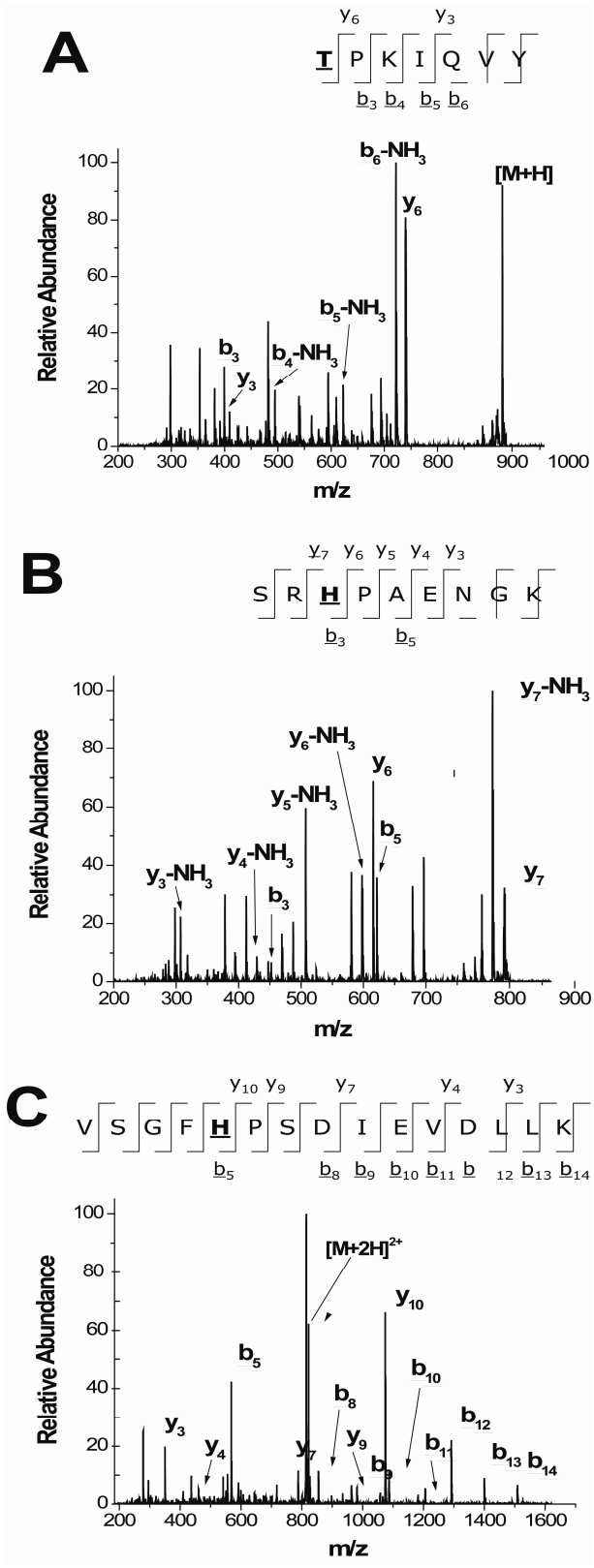
Protein	Residue[a]	% SASA Ratio[b]	SASA of Nε2[b]	pK_a[c]	k[d] ($M^{-1} s^{-1}$)
Cytochrome C	Tyr48	26	30.9	9.8	0.050 ± 0.006
	Tyr67	10	9.0	12.5	0.017 ± 0.002
	Tyr74	27	8.5	10.2	0.015 ± 0.001
	Tyr97	8	7.9	10.9	0
Myoglobin	Tyr103	15	17.1	11.9	0
	Tyr146	4	0	10.1	0
	Ser3	64	3.6	--	0.031 ± 0.003
	Ser117	51	31.5	--	0.028 ± 0.003
β2m	Thr4	63	25.8	--	0.078 ± 0.004
	Ser33	43	25.2	--	0.008 ± 0.002
	Ser88	97	25.7	--	0.021 ± 0.005

[a] All Tyr residues are included for each protein, but only modified Ser and Thr residues are included.

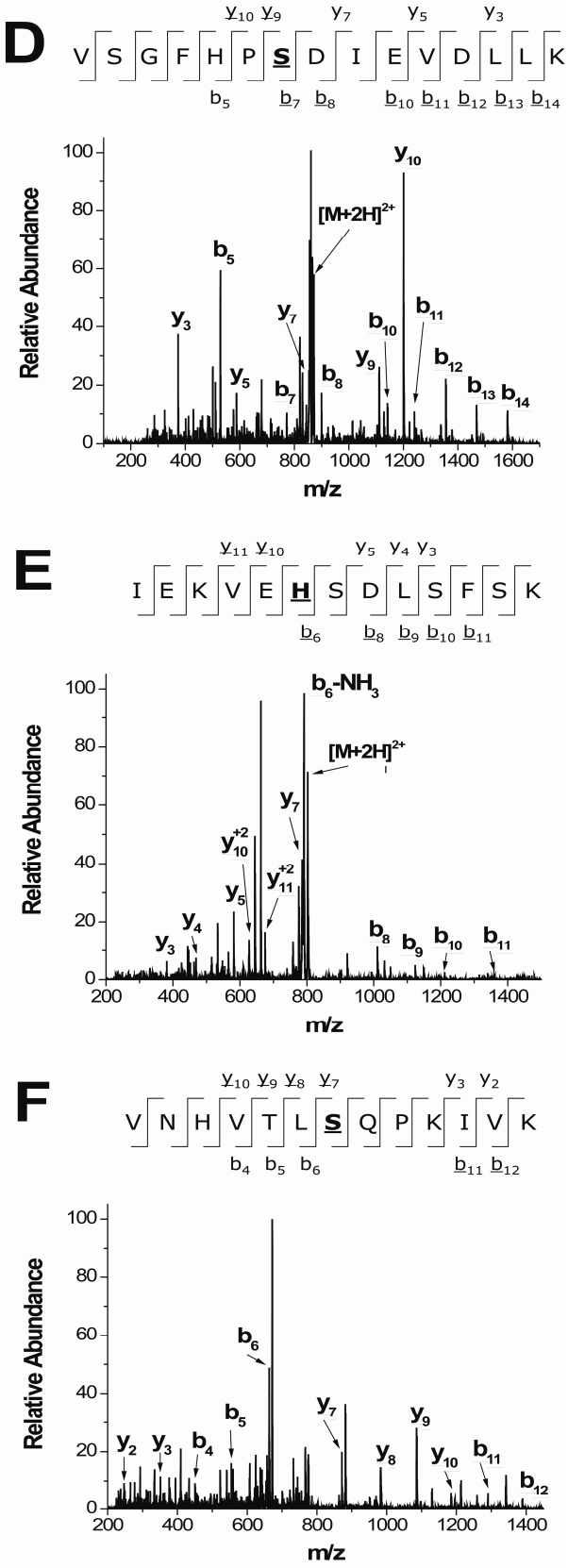
[b] SASA calculated using GETAREA 1.1. 1.4 Å is used as the probe radius, and the calculated SASA is compared to the surface area of the side chain in a Gly-X-Gly tripeptide generating a % ratio. Residues with % ratio below 20 are typically considered buried whereas those with ratios above 30% are regarded as solvent exposed

[c] pK_a calculated using PROPKA

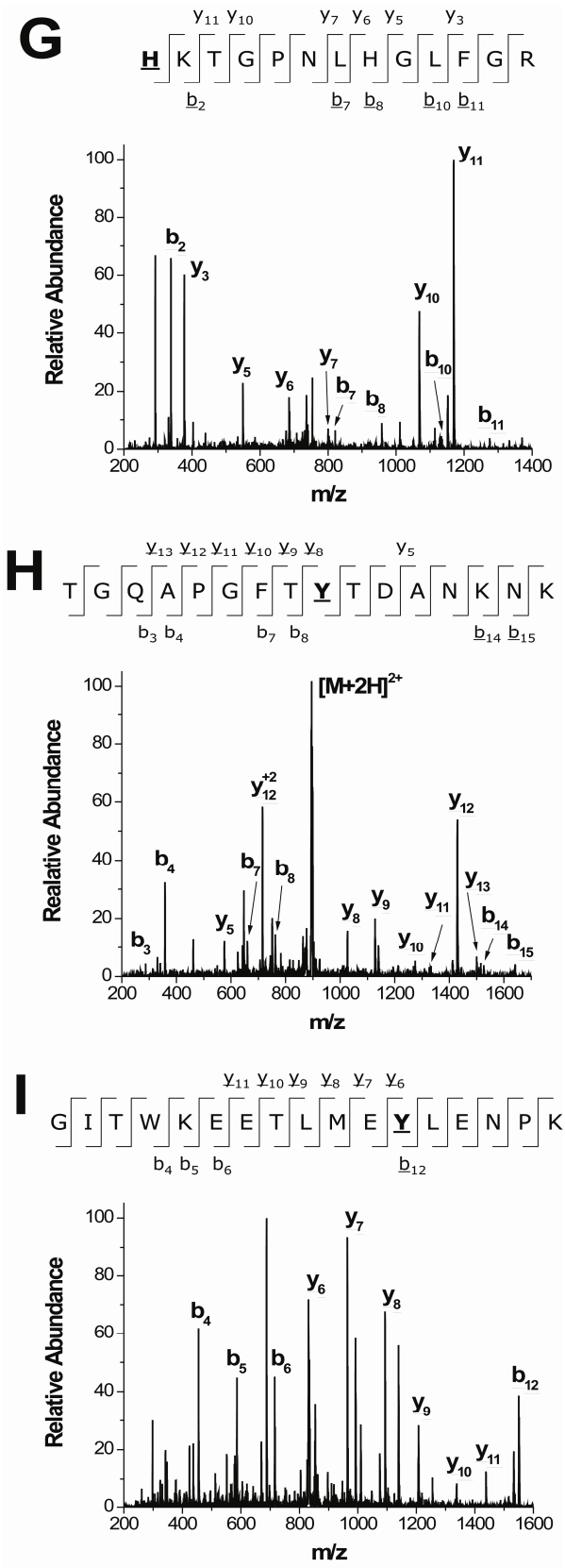
[d] Rate coefficients from dose-response plots like those shown in Figure 2.5



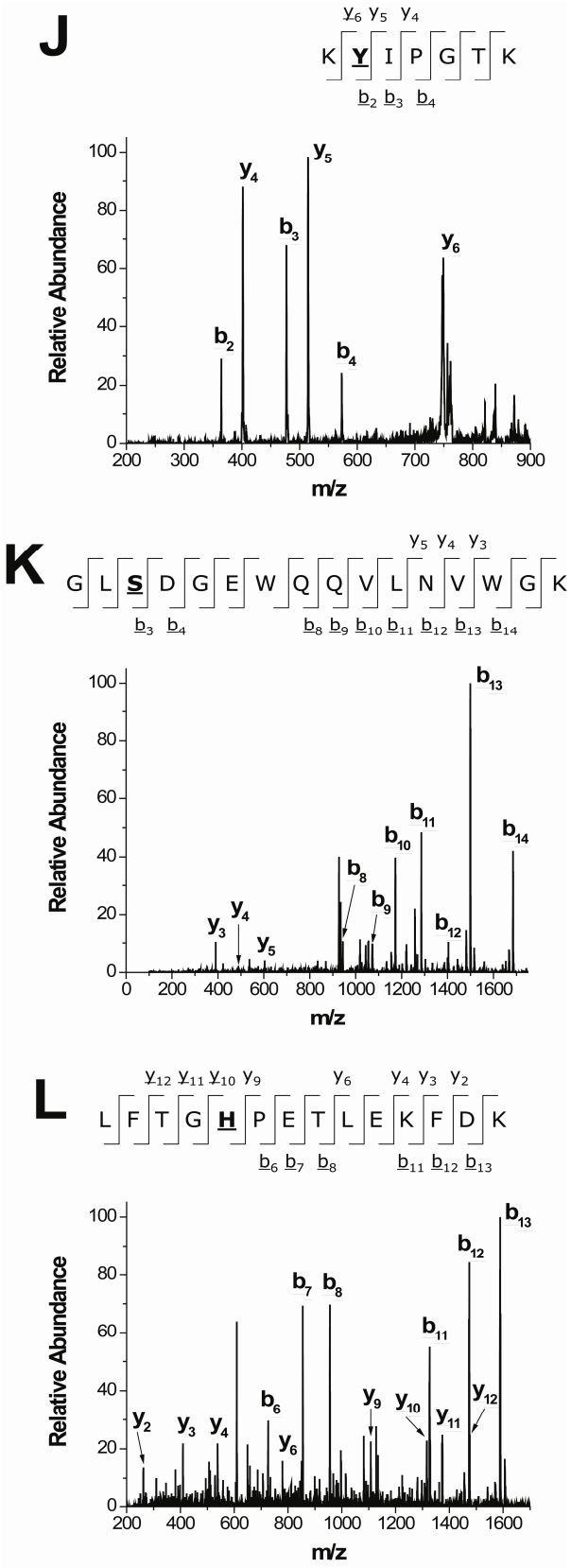
This figure is continued on the next page.



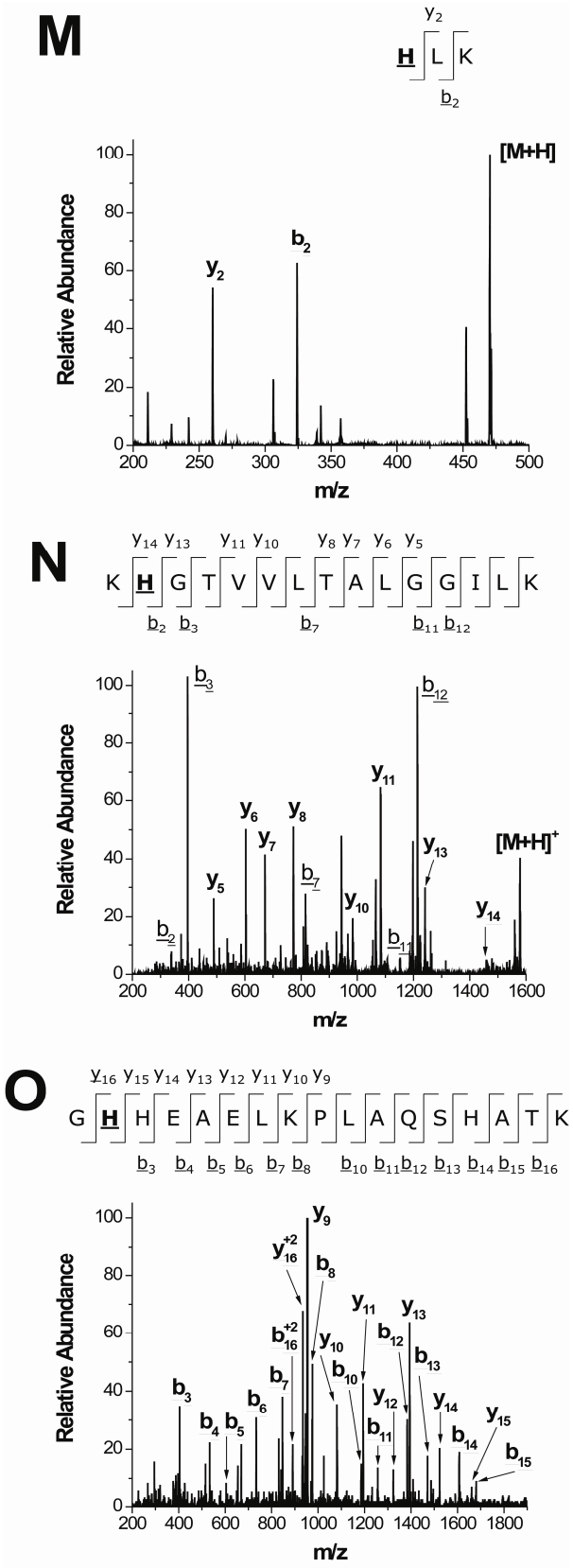
This figure is continued on the next page.



This figure is continued on the next page.



This figure is continued on the next page.



This figure is continued on the next page.

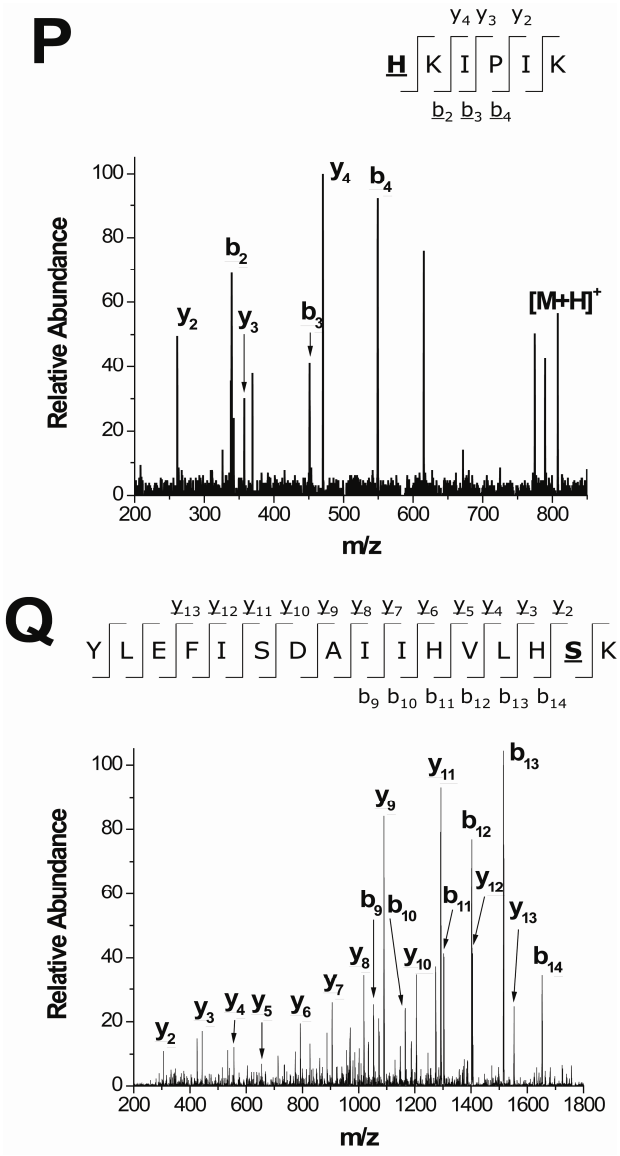


Figure 2.2 MS/MS spectra of β2m (A-F), cytochrome C (G-J), and myoglobin (K-Q) showing modification of: (A) Thr4 (fragment Thr4-Tyr10), (B) His13 (fragment Ser11-Lys19), (C) His31 (fragment Val27-Lys41)*, (D) Ser33 (fragment Val27-Lys41)*, (E) His51 (fragment Val49-Trp60), (F) Ser88 (fragment Val82-Lys94), (G) His26 (fragment His26-Arg38), (H) Tyr48 (fragment Thr40-Lys55), (I) Tyr67 (fragment Gly56-Lys72), (J) Tyr74 (fragment Lys73-Lys79), (K) N-terminus or Ser3 (fragment Gly1-Lys16), (L) His36 (fragment Leu32-Lys45), (M) His48 (fragment His48-Lys50), (N) His64 (fragment Lys63-Lys77), (O) His81 (fragment Gly80-Lys96), (P) His97 (fragment His97-Lys102), and (Q) Ser117 (fragment Tyr103-Lys118)).

The failure to see more modified Ser and Thr residues might be due to two reasons. First, residues that are modified at very low extents can be difficult to detect under normal LC-MS conditions. Second, covalent modification by DEPC is known to be reversible for His and Tyr. For example, the dissociation reaction of modified His has a $t_{1/2}$ of 55 hr at pH = 7 [16]. The modifications of Ser and Thr might also be reversible and could conceivably have shorter $t_{1/2}$, making them even more difficult to detect. To test this second possibility, the reversibility of this reaction was tested for six model peptides that each had at least one Ser, His, or Thr residue (Table 2.3). After reaction of these peptides with DEPC for 1 min, the solutions were stored at 37 ^{o}C for 22 hr to mimic the digestion conditions. The samples then were stored at -10 ^{o}C for the remaining time to mimic the storage conditions used for all the peptides shown in Tables 2.1 and 2.2. Aliquots of these reaction mixtures were analyzed at different time periods. Figure 2.3 clearly shows that the extents of modification of the Ser- and His-containing peptides do not decrease for up to 60 hours, but the extents of modification to the Thr-containing peptides start to decrease after about 10 to 12 hours before leveling off when the samples are frozen. These data suggest that the reaction of DEPC with Thr is reversible and may explain why only a small percentage (~ 5%) of modified Thr residues is detected. Further work is needed to fully characterize the reversibility of this reaction, and methods to avoid this situation will be important in order to expand the practical sequence coverage possible with DEPC.

Table 2.3 Peptides used for investigating reversibility of DEPC modification.

Amino acid modified	Peptide sequence	Peptide mass	Peptide + DEPC masses
Ser	CH_3-GRGD**S**P	601.6	673.6
	LRRA**S**LG	771.9	843.9
His	IARD**H**PYFL	1172	1244
	DRVYI**H**PFHL	1296	1368
Thr	VAI**T**VLVK	842.1	914.1
	GRGD**T**P	601.6	673.6

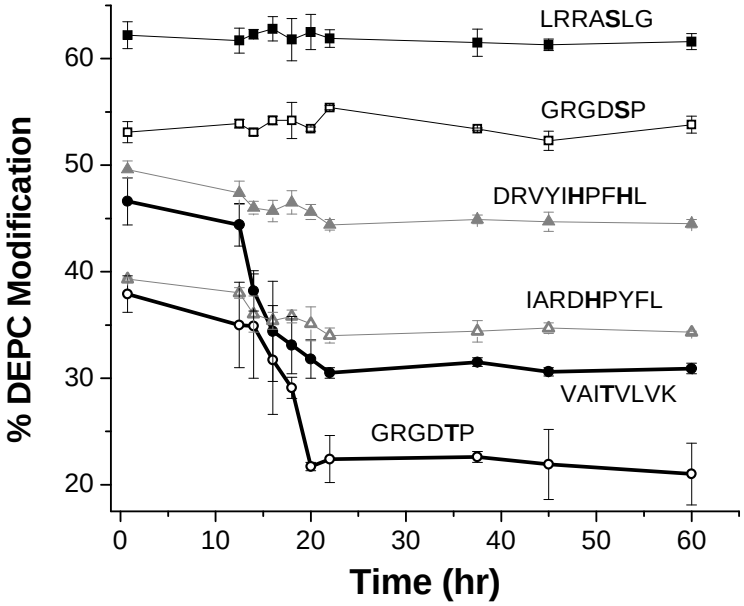


Figure 2.3 Changes in DEPC modification percentages over time for peptides containing Ser, His, and Thr residues. Each plot is labeled with the sequence of the peptide that corresponds to the data. The lines are not mathematical fits of the data but are included to help visualize the qualitative trend. The thicker lines are for the peptides containing Thr.

2.3.2 Dose-Response Curves as Indicators of Changes in Protein Structure

Structural information from surface mapping experiments is reliable only if the protein structure is preserved during the reaction. Some previous surface mapping studies have used spectroscopic methods such as CD or fluorescence spectroscopy to monitor the effect of modification on the protein structure. Some have assumed that as long as the number of modifications is kept low the protein's structure will not be affected by the label. These approaches may make identification of modified residues difficult by limiting the average number of modifications and may lead to errors. CD, for example, provides only the population-weight average properties of a protein sample and may not be sensitive to local structural changes or intra- and intermolecular interactions between protein molecules. Fluorescence spectroscopy can provide local structural information but usually only around Trp residues. Finally, the assumption that limiting the number of modifications will ensure protein integrity may be too simplistic. Some proteins may be able to suffer multiple modifications to surface exposed amino acids without disrupting structure, while the structures of other proteins might be significantly altered by a single modification at a particular site.

Monitoring the extent of protein modification as a function of reagent concentration should be a sensitive method for ensuring that a protein's structure is maintained during the modification reaction. Under the conditions that we use, the reaction of DEPC with a protein is a second-order reaction, and ensuring that this reaction order is maintained at low reagent doses can serve as the means by which protein structural integrity is assured. A second-order reaction is described by equations 2 and 3, where $[P]_o$ is the initial concentration of unmodified protein, $[X]_o$ is the initial

concentration of DEPC, [P] is the unmodified protein concentration at time t, [X] is the DEPC concentration at time t, and k is the second-order rate coefficient.

$$\frac{[P]}{[X]} = \frac{[P]_O}{[X]_O} \; e^{([P]_O - [X]_O)\,kt} \qquad \text{Equation (2)}$$

$$\ln\left(\frac{[X]_O\,[P]}{[X]\;[P]_O}\right) = -\,kt\,[X]_O + kt\,[P]_O \qquad \text{Equation (3)}$$

If a reaction is second order, then a plot of $\ln([P][X]_o)/([P]_o[X])$ vs. $[X]_o$ (or t) will result in a straight line. We assume that the second-order modification rate coefficient will be constant as long as the protein's structure remains unchanged, and such a dose-response plot will thus remain linear. Deviations from linearity or changes in the plot's slope will indicate a change in the reaction dynamics that are likely caused by changes in a residue's microenvironment. Thus, these plots should be very sensitive to any modification-induced structural changes that occur.

The reaction time with DEPC was kept constant at 1 min to minimize DEPC hydrolysis over time, and different DEPC concentrations were used to generate the second-order reaction plots. At low DEPC concentrations, linear relationships are observed between the unmodified protein and the DEPC concentrations, suggesting second-order kinetics and the proteins' 3D structures are maintained during the covalent modification reactions of cytochrome c and myoglobin (Figure 2.4). Deviations from linearity occur at DEPC concentrations above 0.6 mM and 0.4 mM for cytochrome c and myoglobin, respectively, indicating that the proteins' structures are disrupted during 1 min reactions with DEPC above these concentrations. Figure 2.5 shows expanded views of the mass spectra of cytochrome c and myoglobin after reactions with 0.6 mM and 0.4

mM DEPC, respectively. Up to 3 and 4 DEPC adducts are possible for cytochrome c and myoglobin before deviations from linearity are noted. At these DEPC concentrations, the average numbers of modifications per molecule are 1.48 and 0.91 for cytochrome c and myoglobin, respectively. In addition, the modification rate coefficient, k, obtained from the slope of the lines in Figure 2.4 indicates that myoglobin is more reactive with DEPC than cytochrome c. This is most likely due to the greater number of exposed His residues in myoglobin.

While second-order reaction plots of the whole protein data indicate that global changes to protein structure occur above certain DEPC concentrations, an assessment of local changes in protein structure can be obtained by generating second-order reaction plots for individual proteolytic fragments of the proteins. LC-MS analysis of the tryptic peptides of these proteins reveal that the labeling reactions follow second-order kinetics, but there are differences for each fragment both in the concentration at which deviations from linearity are observed and in the modification rate coefficients (Figure 2.6). These observations are important for two reasons. First, deviations from linearity at different DEPC concentrations indicate that certain regions of a protein lose their native structure more readily than others upon reactions with DEPC. Second, the rate coefficients are quantitative indicators of a residue's reactivity and might correlate with a residue's SASA, thus providing more precise structural insight.

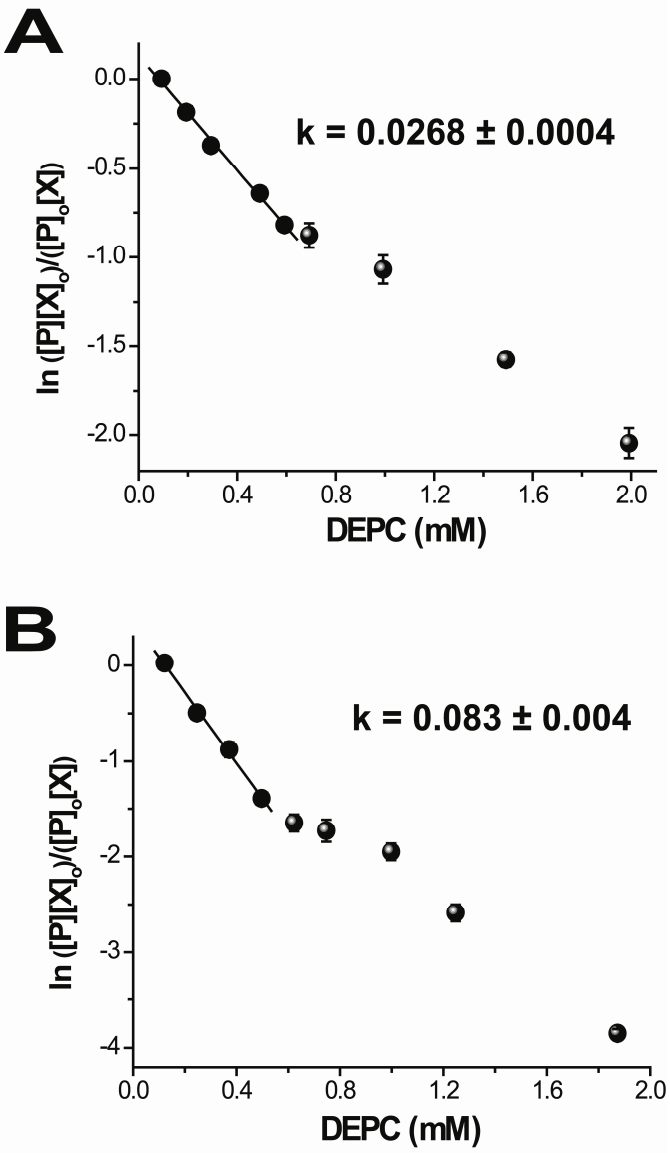


Figure 2.4 Dose-response plots for the reactions of DEPC with (A) cytochrome c and (B) myoglobin. The plot for each protein is produced from the ESI-MS data of the DEPC-treated proteins. The $[P]/[P]_0$ ratio is obtained by dividing the peak area for the unmodified protein by the sum of the peak areas for the modified and unmodified protein. The difference between the [P] and $[P]_o$ is used to determine the concentration of DEPC, [X]. The k values are obtained by dividing the measured slopes by the reaction time.

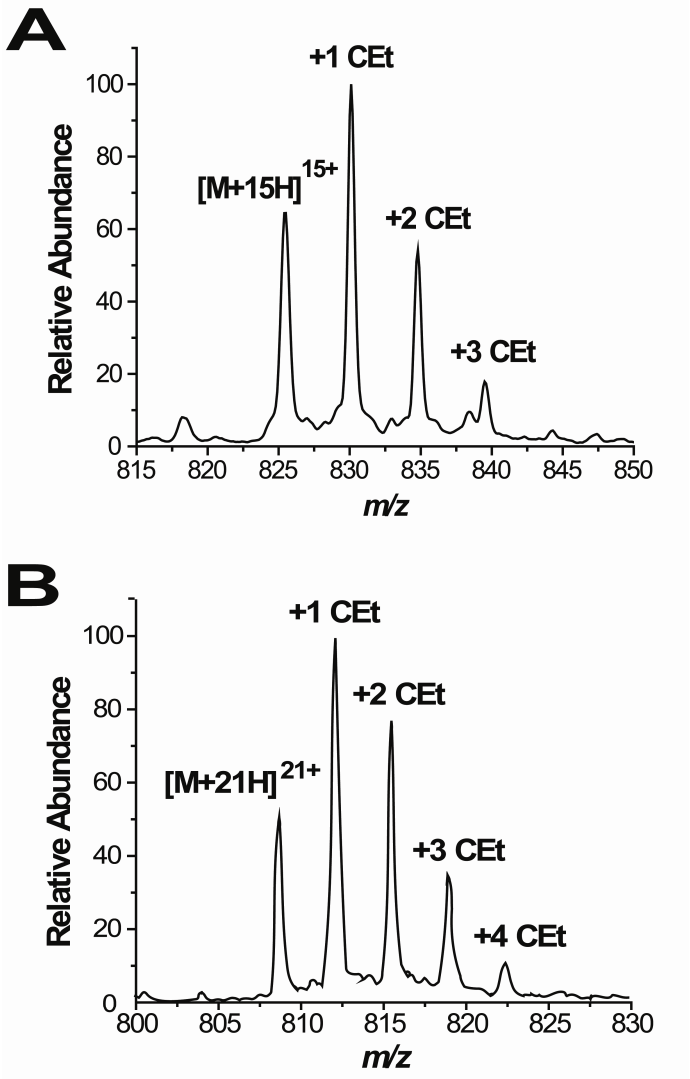


Figure 2.5 Expanded view of mass spectrum showing the extent of DEPC modification for the (A) +15 charge state of cytochrome c and (B) +21 charge state of myoglobin. CEt refers to a carbethoxy group, which is added upon reaction with DEPC.

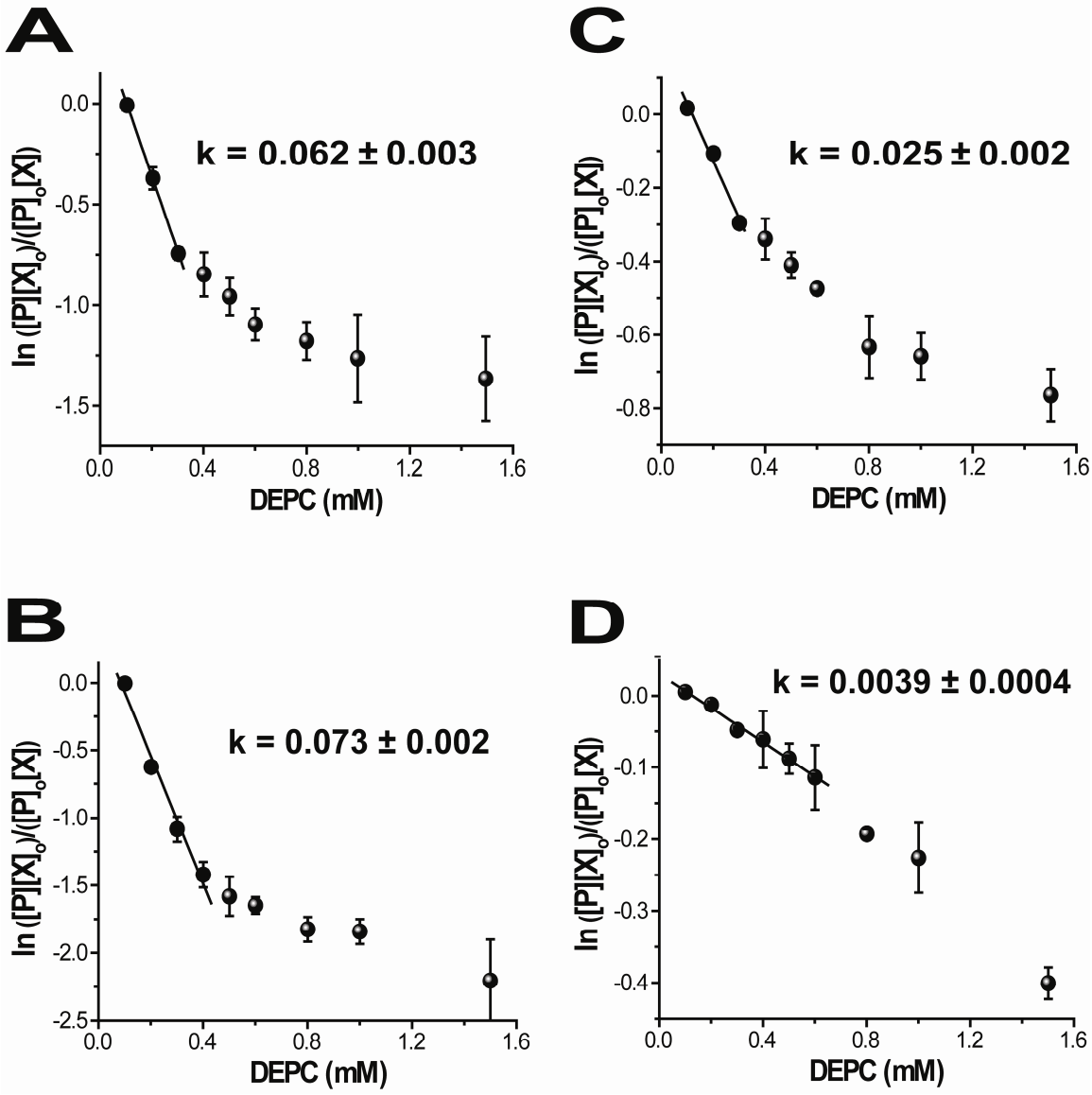


Figure 2.6 Dose-response plots for selected proteolytic fragments of myoglobin after reactions with DEPC. (A) Proteolytic fragment Leu32-Lys45, which contains His36; (B) Proteolytic fragment His48-Lys56; (C) Proteolytic fragment Lys79-Lys96, which contains His81; (D) Proteolytic fragment Gly1-Lys16 which contains Ser3. The plot for each reactive residue is produced from LC-MS data of the proteolytic digests of the modified protein. The ion abundances for the modified and unmodified peptide fragments containing the residue of interest are used to determine [P], $[P]_o$, [X] and $[X]_o$ in a manner similar to that described in the caption of Figure 2.4.

2.3.3 A Comparison with Structural Information from Spectroscopic Techniques

The information obtained from the dose-response plots is compared to CD and fluorescence data in order to (1) check whether the dose-response plots are valid approaches for testing the maintenance of structural integrity and (2) confirm our hypothesis that the dose-response plots are more sensitive to local structural changes in a protein. Tryptophan fluorescence, far-UV CD, and near-UV CD measurements of cytochrome c and myoglobin were made as a function of reagent concentration. Tryptophan fluorescence is commonly used to monitor structural changes in a protein because the fluorescence intensity of this amino acid residue is very sensitive to local structural changes. In this work, the red shift in emission was determined by calculating the average emission fluorescence wavelength, $\langle\lambda\rangle$ [20]. Because it is an integral measurement, the $\langle\lambda\rangle$ is less prone to instrumental noise than is the peak maximum. It is also a more sensitive value because it arises from a calculation involving the entire spectrum, thus, it reflects changes in the shape of the spectrum as well as in position.

Fluorescence measurements (Figure 2.7) of the proteins are consistent with the MS data in Figure 2.4. Changes in the fluorescence intensity of cytochrome c and myoglobin at 1.5 and 0.8 mM DEPC, respectively, suggest that there are structural changes in the proteins at these concentrations. Even though the structural changes monitored by fluorescence occur at higher DEPC concentrations than are observed in Figure 2.4, a comparison of the data from proteolytic fragments containing a Trp residue is most relevant. The DEPC concentration (0.6 mM) at which the fluorescence of myoglobin changes is similar to the concentration at which the linearity of the dose-response plot for the proteolytic fragment Gly1-Lys16 changes (Figure 2.6D). This

peptide fragment contains myoglobin's two Trp residues at positions 7 and 14, and the DEPC-modified Ser residue in this fragment is at position 3. Evidently, the extent of modification that occurs on myoglobin at DEPC concentrations > 0.8 mM is enough to cause a structural change in the N-terminal region of the protein that is reflected in both the fluorescence and MS data. DEPC-modified fragments containing Trp residues were not detected for cytochrome c, so this comparison could not be done for this protein. Even though it is only one comparison, the consistency between the MS and fluorescence data for fragment 1-16 supports the idea that the dose-response plots can act as indicators of protein structural changes. Moreover, the fluorescence data can only provide information about the microenvironment of Trp, which is the least commonly occurring amino acid residue. The dose-response plots from LC-MS analyses, on the other hand, enables monitoring changes in multiple sites in a protein. This appears to be essential for both proteins studied here as the dose-response plots indicate that structural changes occur for some regions of the protein at lower DEPC concentrations than the fluorescence data indicate.

CD spectroscopy is another method that is used to monitor protein structural changes, but this method is much less sensitive to local structural changes. When the far-UV (Figure 2.8) spectra of the proteins in the presence of DEPC are obtained, the results are almost identical to the CD spectra of the native proteins, even for DEPC concentrations up to 2 mM (50-fold excess). For comparison, CD spectra of the proteins in the presence of the denaturant urea are shown to indicate spectral differences that can occur upon changes in protein structure. These results indicate that CD spectroscopy is not as sensitive as the dose-response curves for detecting structural changes that occur

upon covalent modification. Several studies have used CD spectroscopy to check changes in protein structure during covalent labeling experiments [9,14-16,26], but our results suggest that CD spectroscopy may not always be a reliable approach for ensuring protein structural integrity during the modification reactions.

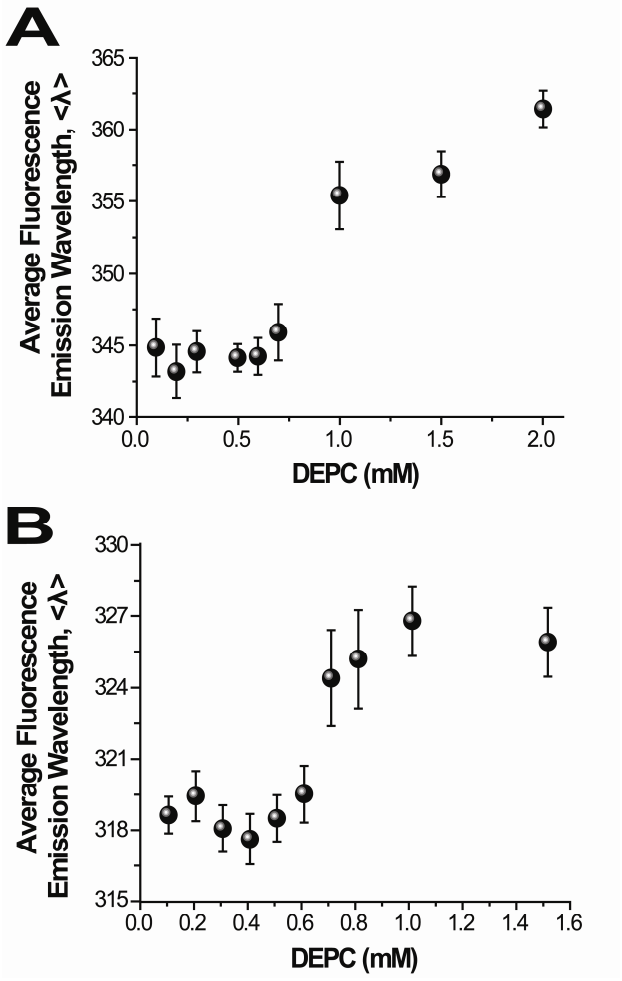


Figure 2.7 Fluorescence measurements of (A) cytochrome c and (B) myoglobin after reactions with different concentrations of DEPC. The fluorescence measurements are reported as average fluorescence emission wavelength calculated as described in Section 2.2.7.

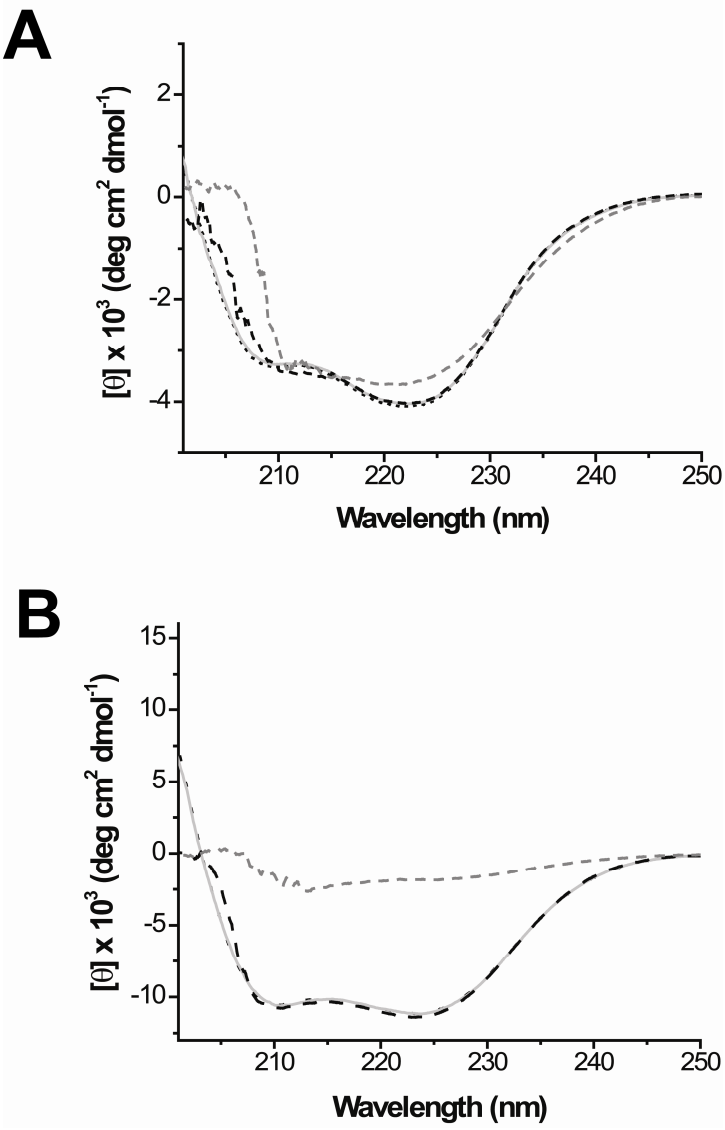


Figure 2.8 Far-UV CD spectra of (A) cytochrome c and (B) myoglobin acquired under native conditions (dotted black line), after a 1 min reaction with a 50-fold excess (5 mM) of DEPC (solid light gray line), in the presence of 1 M urea (dashed black line), and in the presence of 9 M urea (dashed gray line).

2.3.4 Concentration Effects

Because the reaction with DEPC is second-order, the initial protein concentration plays a role in the extent of modification observed. Modification with 10 μM and 100 μM cytochrome c results in different degrees of modification. The average numbers of modifications per cytochrome c molecule are 0.07 and 0.38 for the 10 μM and 100 μM solutions, respectively, at 0.6 mM DEPC. However, we do not expect the rate coefficient to vary with any change in protein concentration. The modification rate coefficients obtained for the two cytochrome c solutions are identical within experimental error with a $k = 0.027 \pm 0.001$ for the 10 μM solution and a $k = 0.0268 \pm 0.0004$ for the 100 μM solution (Figure 2.9). In addition, the range where linearity holds for the 10 μM solution of cytochrome c is higher when compared to the 100 μM solution. In the 10 μM case, the linearity of the plot holds at 10-fold molar excesses of DEPC whereas linearity is lost in the 100 μM solution at molar excesses above 6-fold. These results are not surprising as the higher concentrations of protein and DEPC increase the rate of the reaction, while not changing the rate coefficient for the reaction. A greater number of modifications occur after 1 min in the 100 μM solution, and modification-induced structural changes thus occur at lower DEPC ratios under these conditions. The relevance of these concentration studies has become apparent because of a recent oxidative labeling study [27]. This covalent labeling study using hydroxyl radicals showed that protein concentration has a critical impact on the level of modification, and if not accounted for correctly, data interpretation ambiguities can occur in oxidative labeling studies of protein-protein interactions [27]. Our results clearly show that the second-order plots account for protein concentrations adequately in our covalent labeling studies.

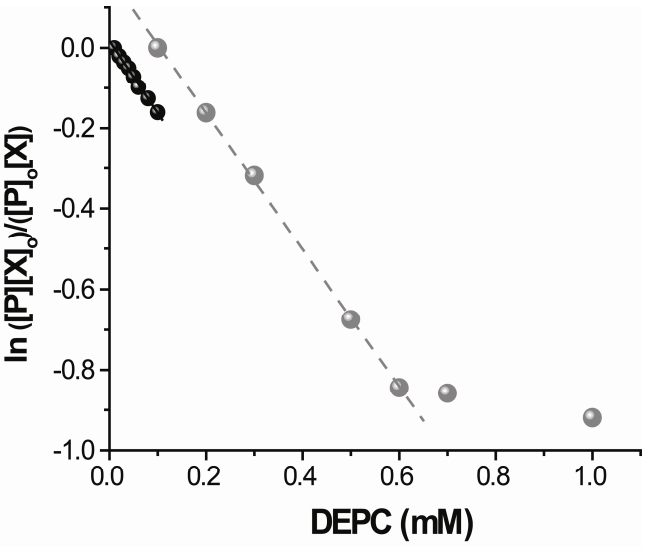


Figure 2.9 Dose-response plot showing the effect of protein concentration on the DEPC modification rate. The modification rate coefficients, k, are 0.027 ± 0.001 for 10 μM (black straight line) and 0.0268 ± 0.0004 for 100 μM (gray dashed line). The k values are obtained by dividing the measured slope by the reaction time.

2.3.5 Factors Affecting the Reactivity of Functional Groups

In addition to providing the range of DEPC concentrations over which a protein's structural integrity is maintained, the dose-response curves provide reaction rate coefficients that allow quantitative comparisons between the reactivity of different residues. Previous studies have shown that reactivity is influenced mainly by the solvent accessibility of the amino acid and of the reactive atom [19,23,28,29]. The rate coefficients in Tables 2.1 and 2.2 indicate that there is not a simple correlation between the modification rates and the solvent accessibility of the entire side chain. The % SASA ratio does not seem to always correlate with the reaction rates. For example, His33 of cytochrome c is not reactive with DEPC despite its high SASA. A better indicator of reactivity appears to be the solvent accessibility of the reactive atom, which is the Nε2

atom in His residues. In cytochrome c, for example, His26 is more reactive with DEPC than His33 despite its lower overall SASA. This is probably due to the higher accessibility of its Nε2 atom. This could also explain why Tyr48 and Tyr74 of cytochrome c do not have the same reactivity with DEPC even though these residues have similar % SASA ratios. Tyr48 is more reactive because its hydroxyl group is more solvent accessible than the hydroxyl group of Tyr74. The lower reactivity of Tyr67, despite a similar OH group SASA as Tyr74, is due to its location in the heme pocket of myoglobin. GETAREA does not calculate SASA with the heme present, so the actual SASA of Tyr67 is surely lower. Similar reasoning explains the lack of reactivity of His93 and His64 of myoglobin and His18 of cytochrome c; the SASA of these residues are artificially high because the heme is not included in the SASA calculation. When the modification rate coefficients for all 18 of the histidine residues from the three proteins in this study are considered, it is clear that only those His residues with Nε2 SASA above 6 $Å^2$ are modified (Tables 2.1) with the exception of 2 His residues. The failure of these two His residues (His93 and His113 from myoglobin) to react with DEPC despite Nε2 SASA above 6 $Å^2$ is actually enlightening about the factors that affect His reactivity. His93 is bound to the heme in myoglobin, so as mentioned above its calculated SASA is incorrect. His113 is probably unreactive because of its close proximity (~ 3.3 Å) to Arg31. This nearby Arg residue might form a H-bond with His113 or its positive charge might hinder the reaction of His113 with DEPC by influencing the partial positive character of the reactive carbonyl carbon of DEPC. The only other His residue in the three proteins that is also close (< 8 Å) to an Arg residue is His31 in β2m. This His residue is reactive, but the modification rate coefficient for this residue is much lower

(0.010 $M^{-1}s^{-1}$) than expected based on its relatively high Nε2 SASA (11.5 $Å^2$). The effect of a nearby charged residue might also be used to explain the lack of reactivity of Tyr103 from myoglobin. Despite its hydroxyl group having a very high SASA, the electrostatic effect of nearby Glu38 (~ 4.3 Å) may lower Tyr103's reactivity. The other Tyr residues do not have charged residues nearby. The relatively few Ser and Thr residues that are reactive make it difficult to assess the factors that affect the reactivity of these residues.

Even stronger quantitative correlations between modification rate coefficients and the accessibility of the reactive atom are difficult to obtain because of the other potential factors that affect reactivity, such as ionization state, H-bonding, and electrostatic effects of nearby residues. Examples of possible electrostatic effects were briefly mentioned above for His113 and Tyr103 of myoglobin and His31 of β2m. A good example of the effect of ionization state is observed upon considering the reactivity of His81 in myoglobin. Despite having the highest Nε2 SASA, the modification rate coefficient of this residue is modest (0.053 $M^{-1}s^{-1}$). Not coincidentally, this His residue also has the highest calculated pK_a of all the His residues, indicating that under the experimental conditions His81 is protonated in most of the protein molecules. Histidine reacts with DEPC only when deprotonated. Because of the many factors that can influence DEPC reactivity, any quantitative description will clearly require the study of a greater number of proteins.

2.4 Conclusions

An improved method of protein surface mapping using covalent modification in conjunction with MS has been described in this chapter. Results demonstrate that DEPC can react with Ser and Thr residues in addition to His and Tyr residues, which dramatically expands the sequence coverage possible with this reagent. Given the prevalence of Ser and Thr residues in proteins, DEPC may be able to compete with non-specific reagents like hydroxyl radicals as a standalone reagent for protein surface mapping. The lesser number of possible products from DEPC modification makes identification of its modification sites more straightforward when compared to hydroxyl radicals that react with many amino acid residues, generating a variety of possible oxidation products. Another advantage of DEPC over hydroxyl radicals is that it does not require the use of expensive synchrotron or other X-ray sources. Further improvements to the method will be necessary to fully access this potential. Improvements such as better detection of modified Ser and Thr residues will expand the practical sequence coverage possible with DEPC.

This work in this chapter also establishes a reliable approach for ensuring protein structural integrity during the modification reactions. Most amino acid-specific covalent labeling studies fail to do this. Dose-response plots are shown to be a useful approach for making sure that the intact structure of a protein is being monitored during the covalent modification reactions. By generating these plots for individual peptide fragments, any small-scale or local perturbations caused by the covalent label can be readily identified and avoided. Furthermore, these plots are more reliable and generally applicable than CD and fluorescence spectroscopies. The dose-response plots provide the added benefit of

quantitative information about residue reactivity, which could possibly provide more precise structural information. While our results cannot currently provide a quantitative description of all the factors that affect reactivity, modification rate coefficients seem to reflect both the reactive atom's solvent accessibility and amino acid microenvironment.

2.5 References

1. Komives, E.A. (2005) Protein-protein interaction dynamics by amide H/H-2 exchange mass spectrometry. *Int. J. Mass Spec. 24*0, 285-290.

2. Eyles, S.J., and Kaltashov, I.A. (2004) Methods to study protein dynamics and folding by mass spectrometry. *Methods 34*, 88-99.

3. Wales, T.E., and Engen, J.R. (2006) Hydrogen exchange mass spectrometry for the analysis of protein dynamics. *Mass Spectrom. Rev.25*, 158-170.

4. Suckau, D., Mak, M., and Przybylski, M. (1992) Protein surface topology-probing by selective chemical modification and mass-spectrometric peptide-mapping. *Proc. Natl. Acad. Sci. USA 89*, 5630-5634.

5. Yem, A.W., Epps, D.E., Mathews, W.R., Guido, D.M., Richard, K.A., Staite, N.D., and Deibel, M.R. (1992) Site-specific chemical modification of interleukin-1 beta by acrylodan at cysteine-8 and lysine-103. *J. Biol. Chem. 267*, 3122-3128.

6. Steiner, R.F., Albaugh, S., Fenselau, C., Murphy, C., and Vestling, M. (1991) A Mass-spectrometry method for mapping the interface topography of interacting proteins, illustrated by the melittin-calmodulin system. *Anal. Biochem. 196*, 120-125.

7. Takamoto, K., and Chance, M.R. (2006) Radiolytic protein footprinting with mass Spectrometry to probe the structure of macromolecular complexes. *Annu. Rev. Biophys. Biomol. Struct. 35*, 251-276.

8. Guan, J.Q., and Chance, M.R. (2005) Structural proteomics of macromolecular assemblies using oxidative footprinting and mass spectrometry. *Trends Biochem. Sci. 30*, 583-592.

9. Dage, J.L., Sun, H., and Halsall H.B. (1998) Determination of diethylpyrocarbonate-modified amino acid residues in alpha(1)-acid glycoprotein by high-performance liquid chromatography electrospray ionization mass spectrometry and matrix-assisted laser desorption/ionization time-of-flight mass spectrometry. *Anal. Biochem. 257*, 176-185.

10. Mailfait, S., Belaiche, D., Kouach, M., Dallery, N., Chavette, P., Formstecher, P., and Sablonnière, B. (2000) Critical role of tyrosine 277 in the ligand-binding and transactivating properties of retinoic acid receptor alpha. *Biochemistry 39*, 2183-2192.

11. Scholten, A., Visser, N.F.C., van den Heuvel, R.H.H., and Heck, A.J.R. (2006) Analysis of protein-protein interaction surfaces using a combination of efficient lysine acetylation and nanoLC-MALDI-MS/MS applied to the E9 : Im9 bacteriotoxin - Immunity protein complex. *J. Am. Soc. Mass Spectrom. 17*, 983-994.

12. Marie, G., Serani, L., Laprévote, O., Cahuzac, B., Guittet, E., and Felenbok, B. (2001) Differential chemical labeling of the AlcR DNA-binding domain from Aspergillas nidulans versus its complex with a 16-mer DNA target: Identification of an essential tryptophan involved in the recognition and the interaction with the nucleic acid. *Protein Sci. 10*, 99-107.

13. Carven, G.J., and Stern, L.J. (2005) Probing the ligand-induced conformational change in HLA-DR1 by selective chemical modification and mass spectrometric mapping. *Biochemistry 44*, 13625-13637.

14. Hassani, O., Mansuelle, P., Cestele, S., Bourdeaux, M., Rochat, H., and Sampieri, F. (1999) Role of lysine and tryptophan residues in the biological activity of toxin VII (Ts gamma) from the scorpion *Tityus serrulatus*. *Eur. J. Biochem. 260*, 76-86.

15. Ehrhardm, B., Misselwitz, R., Welfle, K., Hausdorf, G., Glaser, R.W., Schneider-Mergener, J., and Welfle, H. (1996) Chemical modification of recombinant HIV-1 capsid protein p24 leads to the release of a hidden epitope prior to changes of the overall folding of the protein. *Biochemistry 35*, 9097-9105.

16. Akashi, S., Shirouzu, M., Terada, T., Ito, Y., Yokohama, S., and Takio, K. (1997) Characterization of the structural difference between active and inactive forms of the Ras protein by chemical modification followed by mass spectrometric peptide mapping. *Anal. Biochem. 248*, 15-25.

17. Šantrůček, J., Strohalm, M., Kadlčík, V., Hynek, R., and Kodíček, M. (2004) Tyrosine residues modification studied by MALDI-TOF mass spectrometry. *Biochem. Biophys. Res. Comm. 323*, 1151-1156.

18. Mendoza, V. L., and Vachet, R. W. (2009) Probing protein structure by amino acid-specific covalent labeling and mass spectrometry. *Mass Spec. Rev. 28*, 785-815.

19. Fraczkiewicz, R., and Braun, W. (1998) Exact and Efficient analytical calculation of the accessible surface areas and their gradients for macromolecules. *J. Comp. Chem. 19*, 319-333.

20. Royer, C.A., Mann, C.J., and Matthews, C.R. (1993) Resolution of the fluorescence equilibrium unfolding profile of *trp* aporepressor using single tryptophan mutants. *Protein Sci. 2*, 1844-1852.

21. Glocker, M.O., Kalkum, M., Yamamoto, R., and Schreurs, J. (1996) Selective biochemical modification of functional residues in recombinant human macrophage colony-stimulating factor beta (rhM-CSF beta): Identification by mass spectrometry. *Biochemistry 35*, 14625-14633.

22. Tsubaki, M., Kobayashi, K., Ichise, T., Takeuchi, F., and Tagawa, S. (2000) Diethyl pyrocarbonate modification abolishes fast electron accepting ability of cytochrome b(561) from ascorbate but does not influence electron donation to monodehydroascorbate radical: Identification of the modification sites by mass spectrometric analysis. *Biochemistry 39*, 3276-3284.

23. Kalkum, M., Przybylski, M., and Glocker, M.O. (1998) Structure characterization of functional histidine residues and carbethoxylated derivatives in peptides and proteins by mass spectrometry. *Bioconjugate Chem. 9*, 226-235.

24. Lundblad, R. L., and Noyes, C. M. (1984) *Chemical Reagents for Protein Modification*, pp. 105-126, CRC Press, Boca Raton, FL.

25. Trinquier, G., and Sanejouand, Y.H. (1998) Which effective property of amino acids is best preserved by the genetic code? *Protein Eng. 11*, 153-169.

26. Gao, J., Zhang, F., Zhang, L., Guo, Y., Ruan, K., Jiang, D., and Xu, C. (2007) Six specific lysine residues are crucial in maintaining the structure and function of soluble manganese stabilizing protein. *Acta Biochim. et Biophys. Sinica* **38**, 611-619.

27. Tong, X., Wren, J.C., and Konermann, L. (2007) Effects of protein concentration on the extent of gamma-ray-mediated oxidative Labeling studied by electrospray mass spectrometry. *Anal. Chem. 79*, 6376-6382.

28. Strohalm, M., Šantrůček, J., Hynek, R., and Kodíček, M. (2004) Analysis of tryptophan surface accessibility in proteins by MALDI-TOF mass spectrometry. *Biochem. Biophys. Res. Comm. 323*, 1134-1138.

29. Novak, P., Kruppa, G.H., Young, M.M., and Schoeniger, J. (2004) A top-down method for the determination of residue-specific solvent accessibility in proteins. *J. Mass Spectrom. 39*, 322-328

CHAPTER 3

COPPER BINDING AND BINDING-INDUCED CONFORMATIONAL CHANGES IN β-2-MICROGLOBULIN

This chapter is part of papers published as: Srikanth, R., Mendoza, V. L., Bridgewater, J. D., Zhang, G., and Vachet, R. W. (2009) Copper binding to β-2-microglobulin and its pre-amyloid oligomers. *Biochemistry 48*, 9871-9881 and Mendoza, V.L., Antwi, K., Baron-Rodriguez, M.A., Blanco, C., and Vachet, R.W. (2009) Structure of the pre-amyloid dimer of β-2-microglobulin from covalent labeling and mass spectrometry. *Biochemistry 49*, 1522-1532.

3.1 Introduction

The cause of β-2-microglobulin (β2m) amyloid formation *in vivo* is not precisely known, but several means of causing fibril formation *in vitro* have been identified. β2m amyloid fibrils can be generated under acidic conditions (pH < 3.6) [1], by removing the first six N-terminal amino acids [2], by mixing the protein with collagen at pH = 6.4 [3], by sonicating the protein in the presence of sodium dodecyl sulfate at pH = 7.0 [4], and by incubating the protein at physiological conditions in the presence of stoichiometric amounts of Cu(II) [5,6].

The latter means of initiating β2m fibril formation is interesting for a number of reasons. Several proteins are known to form amyloid fibrils in the presence of divalent metals, especially Cu(II), thus, Cu(II)-protein interactions seem to represent one of the general motifs for stimulating amyloid formation [7-11]. For β2m, the case has been made that Cu(II) could be the initiating factor *in vivo* because of the elevated Cu(II) concentrations that hemodialysis patients experience (i.e. dialysate and dialysis membrane) [5,12]. The conditions necessary to stimulate β2m fibril formation in the presence of Cu(II) are also arguably more similar to physiological conditions than other

methods used to stimulate β2m fibril formation *in vitro*. In addition, a catalytic role for Cu(II) is suggested by recent studies that found Cu binding to be essential for initiating oligomer formation under near-physiological conditions (i.e. pH 7.4, 37 ºC, 150 mM ionic strength) *in vitro* but not essential for maintaining the stability of tetramers or hexamers and completely absent in the eventual amyloid fibrils [13,14]. While an *in vivo* role for Cu(II) in β2m amyloid formation is not confirmed, the metal does represent a discrete way to trigger amyloid formation so that the intermediates that precede the fibrils can be more easily studied. The added appeal of studying Cu(II) as a trigger is that the structural changes brought about by Cu binding are subtle and the pre-fibril oligomers have native-like structures [15]. Intriguingly, very similar divalent metals, such as Ni(II), do not cause β2m to self assemble [15], indicating that the mode of Cu(II) binding is unique enough to cause just the right structural changes to facilitate oligomerization and fibril formation.

Given the importance of β2m fibril formation and the unique ways in which Cu(II) catalyzes this reaction, the structural details of how Cu(II) binds the monomer and causes the necessary structural changes for oligomerization are important to determine. In this chapter, covalent labeling with mass spectrometric detection is used to determine the Cu(II) binding site in β2m and any conformational changes induced by metal binding. These covalent labeling methods provide information about protein structure and interactions by comparing the reactivity of amino acid side chains in solution under different conditions (i.e. with and without Cu). The covalent labeling reactions can be used to identify Cu(II) binding sites because reactivity with these reagents is significantly decreased when residues are bound to Cu [16,17].

3.2 Experimental Procedure

3.2.1 Materials

Human β2m was obtained from Fitzgerald Industries International, Inc. (Concord, MA). Diethylpyrocarbonate (DEPC), 2,3-butandeione (BD), sulfo-*N*-hydroxysuccinimide acetate (NHSA), imidazole, dithiothreitol (DTT), copper-(II) sulfate (CuSO4), 3-morpholinopropanesulfonic acid (MOPS), potassium acetate, arginine, ubiquitin from bovine erythrocytes, equine heart cytochrome c, equine skeletal muscle myoglobin, chicken egg white ovalbumin, human hemoglobin, and bovine transferrin were purchased from Sigma-Aldrich (St. Louis, MO). Tris(hydroxymethyl)-aminomethane (Tris) was purchased from EM Science (Gladstone, NJ). Urea was purchased from Mallinckrodt Chemicals (Phillipsburg, NJ). Trypsin was from Promega (Madison, WI), and chymotrypsin was purchased from Roche Diagnostics (Indianapolis, IN). Centricon molecular weight cutoff (MWCO) filters were obtained from Millipore (Burlington, MA). Deionized water was prepared from a Millipore (Burlington, MA) Simplicity 185 water purification system.

3.2.2 β2m Solutions

100 μM β2m in 200 mM potassium acetate, 500 mM urea, and 25 mM MOPS (pH 7.4) with 200 μM copper(II) sulfate at 37 °C. All components were equilibrated at 37 °C prior to Cu(II) addition, and immediately returned to 37 °C after mixing. The covalent labels were added to the incubated β2m 2 minutes after addition of Cu(II).

3.2.3 Carbethoxylation with DEPC

Stock solutions of DEPC were prepared in acetonitrile. The DEPC reactions of β2m were performed for 1 min at 37 °C and were initiated by adding 0.25 mM DEPC. The total reaction volume for the experiments was 30 μL, and the total amount of acetonitrile added was ~ 1.5 %. The reactions were quenched after 1 min by adding 10 mM imidazole.

3.2.4 Acetylation with NHSA

Stock solutions of NHSA were prepared in water. The labeling of β2m with 0.30 mM NHSA was carried out for 1 min at 37 °C. The total reaction volume for the experiments was 30 μL. The reactions were quenched by adding 10 mM Tris.

3.2.5 BD Modification

Stock solutions of BD were prepared in water. The reactions of β2m with 35 mM BD were performed in the dark for 1 min at 37 °C. Reactions were carried out in the dark to avoid possible photoactivation of the label, which could enhance nonspecific reactions with residues other than arginine [18,19]. The total reaction volume for the experiments was 30 μL. The reactions were quenched by adding 100 mM arginine.

3.2.6 Proteolytic Digestion

Before proteolytic digestion, all modified samples were purified using a 10,000 MWCO filter and reconstituted with deionized water to a final concentration of 300 μM. Purified β2m samples in 25 mM Tris-HCl (pH 7) and 1 mM $CaCl_2$ were first reacted with

10 mM DTT at 37 °C for 45 min to reduce the disulfide bonds. This was followed by addition of acetonitrile and incubation at 37 °C for 45 min. Trypsin and chymotrypsin (1 μg/μL) were then added to DEPC-modified samples to yield a final enzyme/substrate ratio of 1:20. For the NHSA- and BD-modified β2m samples, only chymotrypsin (1 μg/μL) was added. Although trypsin is the most reliable and robust protease, chymotrypsin was used instead because trypsin can no longer cleave proteins after acetylated lysine and modified arginine residues. All samples were digested at 37 °C for 16 h. The enzymes were inactivated by adding 2 μL of acetic acid, and the samples were immediately analyzed.

3.2.7 Instrumentation

The amount of modification was determined by removing an aliquot of the purified β2m and analyzing the samples using a Bruker Esquire-LC quadrupole ion trap mass spectrometer (Billerica, MA) equipped with an ESI source. The ESI source was operated at a spray voltage of 3.5 kV, and the capillary temperature was set at 300°C. The voltages for the transfer optics between the ESI source and the ion trap were optimized for maximum signal, with typical skimmer 1 and capillary offset values of 30-35 V and 50-60 V, respectively.

The proteolytic fragments were separated by an Agilent HP1100 (Wilmington, DE) HPLC system with a C18 column (15 cm x 2.1 mm, 5 μm particle size, Supelco, St. Louis, MO) for on-line analysis by MS and MS/MS. The fragments of NHSA- and BD-modified β2m were eluted using a linear gradient of methanol that increased from 5 to 70% over 20 min and 70 to 100% over the final 2 min at a flow rate of 0.250 mL/min.

For the fragments of DEPC-modified β2m, a linear gradient of methanol that increased from 5 to 70% over 30 min and 70 to 100% over the final 3 min was used. For both gradient conditions, water comprised the balance of the solvent, and a total of 0.1% acetic acid was present. The LC effluent was fed into the mass spectrometer with similar ESI source conditions as described above. Tandem mass spectra were acquired using collision induced dissociation (CID) with isolation widths of 1.0 Da and excitation voltages between 0.6 and 1.0 V. Peptide sequences were determined from the MS/MS data via *de novo* sequencing or with the help of BioTools (Bruker Daltonics, Billerica, MA).

Sequencing of β2m to identify the specific amino acids modified by BD is not possible using CID, so electron transfer dissociation (ETD) was used to identify the specific modification sites. In the CID process the BD label is readily lost without any associated information about its location in the sequence. The ETD experiments were carried out using a Bruker HCT ETD II (Billerica, MA) equipped with an ESI source. ETD was performed using total reaction time of 25 ms and low mass cutoff of 100. Accumulation times for the fluoranthene radical anion, which was used as the ETD reagent, were typically between 3 and 5 ms.

3.2.8 Amino Acid Modification Percentage

The percent modification of each labeled amino acid was determined by comparing the LC-MS intensities of modified and unmodified proteolytic peptide fragments containing the amino acid of interest. For each peptide fragment, the modified form eluted after the unmodified form with the difference in retention times ranging from 1-12 minutes. The ion intensities of both modified and unmodified peptides were

determined from extracted ion chromatograms (EIC). As an example, Figure 3.1 shows the EIC of the unmodified (m/z 623.4) and modified (m/z 644.4 and 665.4) forms of the fragment Ile1-Tyr10 in the absence of Cu(II). The EIC for all peptides is averaged over a 0.4 min time range to determine the ion intensities. The percent modification was obtained by dividing the ion intensity of the modified fragment ($I_{modified}$) by the sum of the ion intensities for the modified ($I_{modified}$) and unmodified ($I_{unmodified}$) fragments as shown in equation 1. The errors are reported as the standard error of the mean using the data from three separate modification reactions. Because ion intensity ratios were used, low modification levels could be accurately determined.

$$\% \text{ Modification} = \frac{I_{\text{modified}}}{I_{\text{modified}} + I_{\text{unmodified}}} \times 100 \qquad \text{Equation (1)}$$

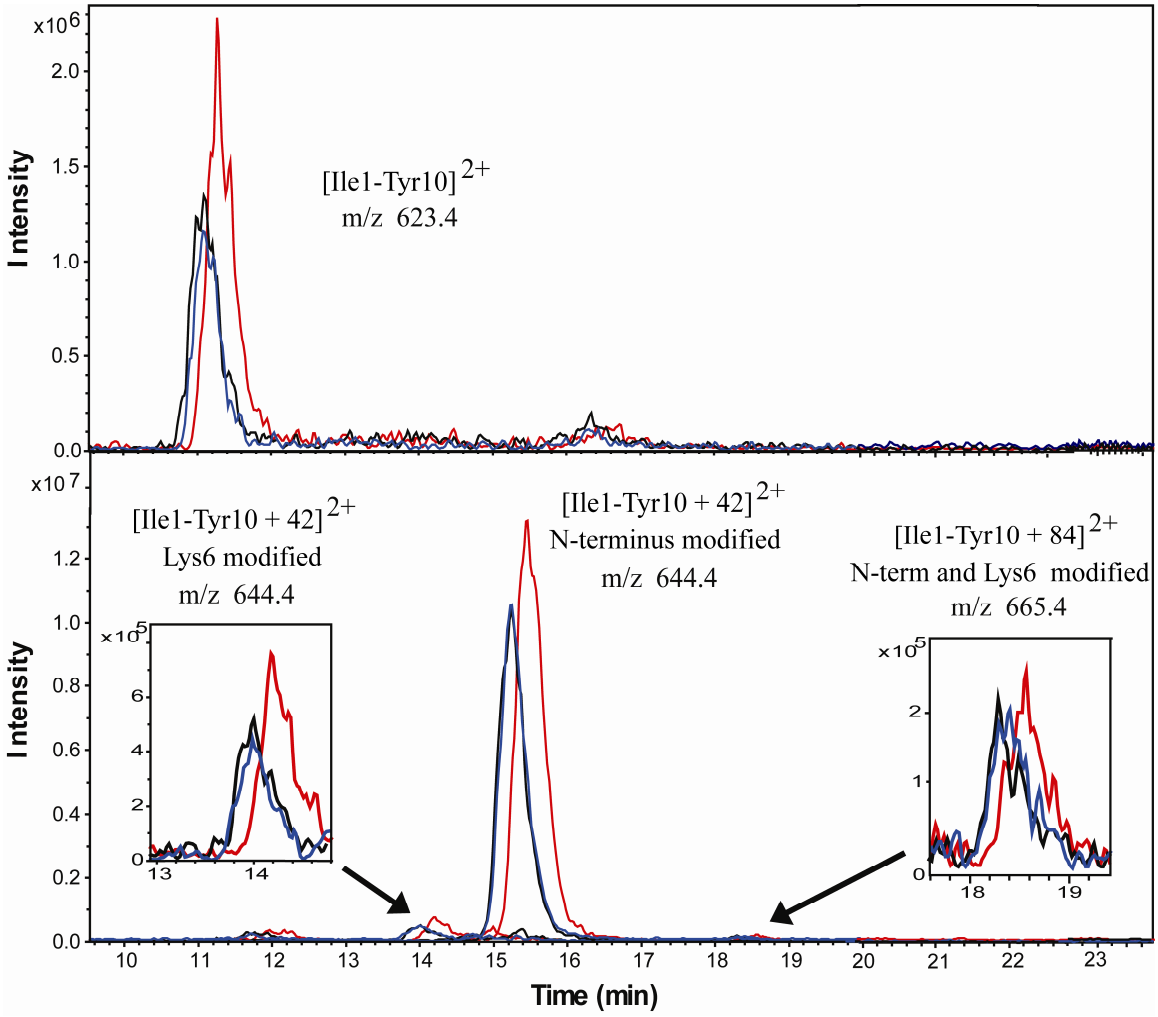


Figure 3.1 Extracted ion chromatograms (EIC)of the unmodified (m/z 623.4) and modified (m/z 644.4 and 665.4) forms of the fragment Ile1-Tyr10 in the absence of Cu(II). Each trace represents one trial. The peaks at ~14 min and ~15 min are the Lys6- and N-terminus-modified fragments, respectively. The peaks at ~18.2 min are the di-modified fragments.

3.3 Results and Discussion

The reactivity of monomeric β2m in the absence and presence of Cu(II) was monitored using three covalent labeling reagents – diethylpyrocarbonate (DEPC), sulfo-*N*-hydroxysuccinimide acetate (NHSA), and butanedione (BD). DEPC can react with solvent exposed His, Lys, Tyr, Thr, and Ser residues [20], but this reaction is significantly slowed when these residues, especially His, are bound to Cu(II) [16,17]. Likewise, NHSA and BD can react with amino and guanidinium groups, respectively, when these sites are solvent exposed [20], but its reactivity is reduced when these residues are protected, for example, by metal binding. Moreover, amino acid-specific information about β2m's structural changes can be inferred from differential reactivity patterns of individual amino acids.

3.3.1 Covalent Labeling with NHSA

NHSA can react with amino groups such as the ε-NH_2 of lysine residues and the N-terminal α-NH_2. First, the optimum reaction time and molar ratio of NHSA were determined. As described previously [20,21] and in chapter 2, plots of the 2^{nd} order reaction kinetics for the modified peptide fragments can be used to ensure that the covalent label itself does not disrupt the structure of β2m near a particular reactive amino acid residue (Figures 3.2A and 3.2B). For monomeric β2m, deviations from linearity, which are indicators of label-induced structural perturbations, occur at a 4-fold molar excess of NHSA for most of the modified proteolytic peptide fragments. Hence, a 3-fold molar excess of the label was added to the β2m samples to avoid label-induced protein

structural changes. Upon modification with 3-fold excess of NHSA for 1 min at 37 ^{o}C, addition of up to three acetyl groups is reproducibly observed, with the average number of modifications to the whole protein around one (Figure 3.2C).

Identification of the modified amino acids was achieved by proteolytic digestion and LC-MS/MS analyses. LC-MS/MS analyses indicate that the N-terminus, almost all of the lysines (Lys6, Lys19, Lys41, Lys58, Lys75, Lys91, and Lys94), and Asn83 are labeled to different degrees. The unmodified and modified fragments containing the lysine residues were detectable at t=0 and t=2min after addition of Cu(II). Modification of an Asn residue, Asn83, is unprecedented as NHSA has only been reported to react with N-terminal amines and lysine residues [22], but the tandem mass spectrum of the modified fragment Ala79-Leu87 unambiguously identifies Asn83 as the modified amino acid.

When the covalent labeling reactions with NHSA are performed in the absence and presence of Cu(II), the modification extents of some residues change (Table 3.1). Because the percent modification is determined using the ion intensity ratios of the unmodified and modified peptide fragments, small changes in modification levels can be accurately and precisely determined. From the data in Table 3.1, we find that the levels of modification of all but two of the residues remain approximately the same after the addition of Cu(II). The reactivity of the N-terminus and Lys91 decrease by 25% and 18%, respectively, after addition of Cu(II).

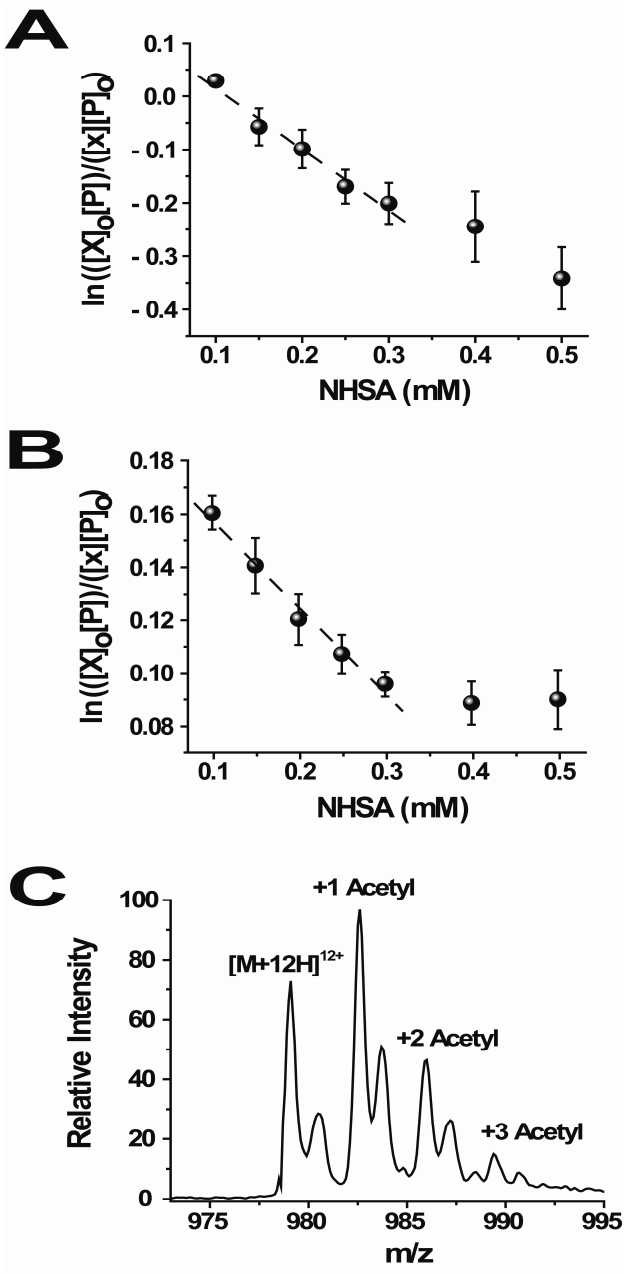


Figure 3.2 (A) Dose-response plot for Lys6 (fragment Thr4-Tyr10) after reaction with NHSA. (B) Dose-response plot for Lys75 (fragment Tyr67-Tyr78) after reaction with NHSA. The plot for each reactive residue is produced from LC-MS data of the proteolytic digests of the modified protein. The $[P]/[P]_o$ ratio is obtained by dividing the peak area for the unmodified fragment by the sum of the peak areas for the modified and unmodified fragments. The difference between the [P] and $[P]_o$ values is used to determine the concentration of NHSA, [X]. (C) Expanded view of mass spectrum showing the extent of acetylation for the +12 charge of β2m at t=0. The second peak of each doublet is due to a fraction of the protein being oxidized at Met99. This residue is easily oxidized during protein storage. The portion of the protein molecules that are oxidized is typically about 20 to 30% of the total protein.

3.3.2 Covalent Labeling with DEPC

DEPC reacts readily with histidine residues but can also react with amine and hydroxyl groups at neutral pH [20,21 and chapter 2]. The DEPC experiments were performed in much the same way as the lysine acetylation studies. Based on 2nd order kinetics plots of each modified peptide, a 2.5-fold molar excess of DEPC was used to minimize DEPC-induced β2m structural changes (Figure 3.3A and 3.3B). As was the case with the NHSA reactions, the reaction time was kept short (1 min) to avoid modification-induced structural change and to minimize DEPC hydrolysis by water. Under the experimental conditions used, the addition of up to four carbethoxy (CEt) groups to the protein was observed, and the average number of modification was between one and two (Figure 3.3C). Proteolytic digestion of the protein and LC-MS/MS analyses indicate that DEPC reacted with the N-terminus, Thr4, Lys6, His13, Lys19, Tyr26, Ser28, His31, Ser33, Lys41, His51, Ser57/Lys58, Tyr63, Tyr67, Lys75, Ser88, and Lys94. These modifications were observed at t=0 and t=2min after addition of Cu(II), but the level of modification for some residues changed over time (*vide infra*). In all cases but one (Ser57/Lys58), the specific amino acids that were modified could be determined unambiguously from the MS/MS data.

The reactions with DEPC reveal that some residues change in reactivity upon addition of Cu(II) (Table 3.1). LC-MS data reveal that the percent modifications of six residues decrease after Cu(II) binding. These residues are the N-terminus, Ser28, His31, Ser33, Tyr63, and Tyr67. His31 undergoes the most significant change in reactivity, with about a 61% decrease in modification when Cu(II) is present. A decrease in extent of modification between 30 and 40% was observed for the N-terminus, Ser28, Ser33, and

Tyr63. Tyr67 showed the least drop in reactivity, ~17% in the presence of Cu(II). The changes in reactivity observed for the N-terminus, Lys6, Lys19, Lys41, Lys58, Lys75, and Lys94 with DEPC agree with the NHSA data described above.

3.3.3 Covalent Labeling with Butanedione

Unlike DEPC, which reacts with many nucleophilic groups, BD reacts specifically with arginine. The reaction is reversible at pH < 9, so higher reagent doses are necessary to improve the product yield. The excess reagent allows pseudo 1st order kinetic plots to be used to ensure the structural integrity of the protein (Figure 3.4A). Based on these plots, 350 fold molar excess (i.e. 35 mM) was chosen. Under our reaction conditions, up to two modifications were detected with an average number of modification of less than one (Figure 3.4B). The LC/MS and ETD data indicate that BD reacts with Arg3, Arg12, Arg45, and Arg97 but not Arg81.

The reactions of BD with β2m reveal that three of the four reacting arginine residues undergo changes in reactivity after the addition of Cu(II) (Table 3.1). The reactivities of Arg3 and Arg97 dramatically increase (about >70%) immediately after Cu(II) is added. In contrast, the level of modification of Arg12 decreased by 23%.

A summary of the extent of modification for each of the modified amino acids in the absence and presence of Cu(II) is shown in Table 3.1.

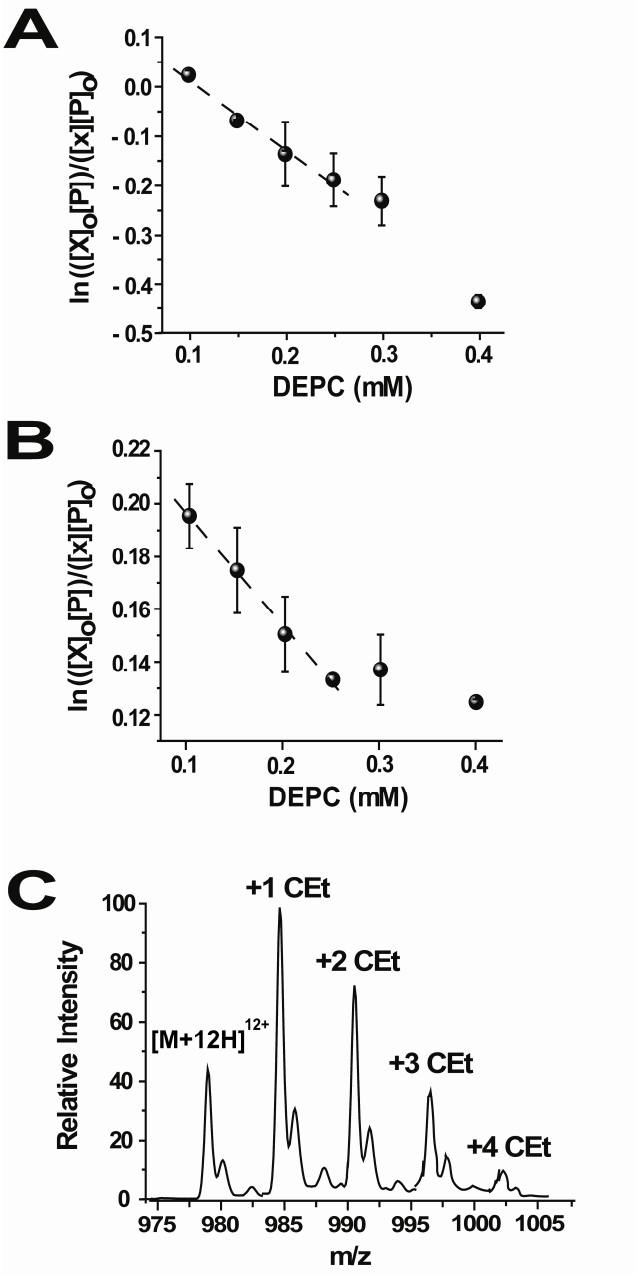


Figure 3.3 (A) Dose-response plot for His13 after reaction with DEPC. (B) Dose-response plot for Tyr26 after reaction with DEPC. The plot for each reactive residue is produced from LC-MS data of the proteolytic digests of the modified protein. The $[P]/[P]_o$ ratio is obtained by dividing the peak area for the unmodified fragment by the sum of the peak areas for the modified and unmodified fragments. The difference between the [P] and $[P]_o$ values is used to determine the concentration of DEPC, [X]. (C) Expanded view of mass spectrum showing the extent of carbethoxylation for the +12 charge of β2m at t=0. The second peak of each doublet is due to a fraction of the protein being oxidized at Met99. This residue is easily oxidized during protein storage. The portion of the protein molecules that are oxidized is typically about 20 to 30% of the total protein.

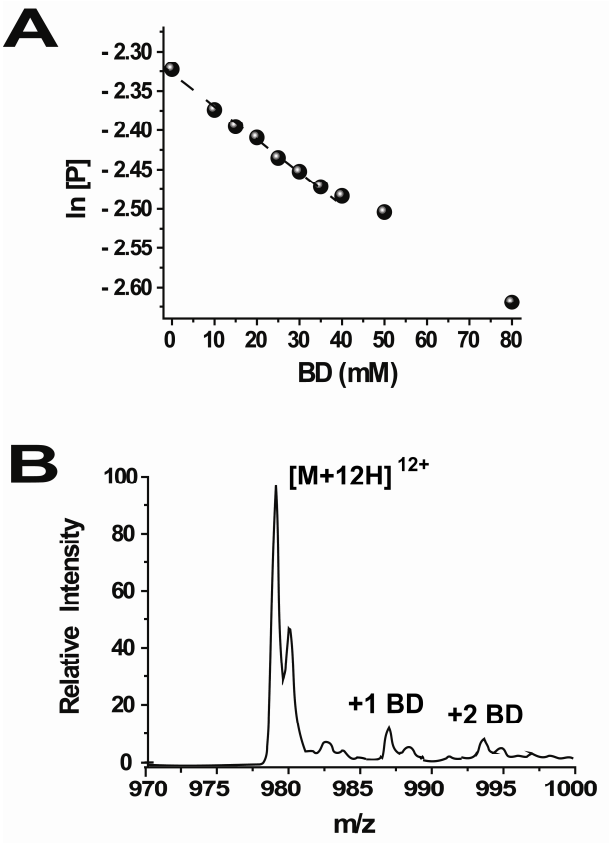


Figure 3.4 (A) Dose-response plot for β2m after reaction with BD. The plot is produced from the ESI-MS data of the BD-treated proteins. The [P] is the remaining concentration of β2m after reaction with BD. (B) Expanded view of mass spectrum showing the extent of modification for the +12 charge of β2m at t=0. The second peak of each doublet is due to a fraction of the protein being oxidized at Met99. This residue is easily oxidized during protein storage. The portion of the protein molecules that are oxidized is typically about 20 to 30% of the total protein.

Table 3.1 Summary of the percentage changes for the modified amino acids of β2m in the absence and presence of Cu(II).

Residue	apo-β2m	Cu(II)	% Change
NHSA			
N-terminus	81 ± 3	60.5 ± 0.7	- 29
Lys6	20 ± 1	19 ± 1	0
Lys19	7.0 ± 0.4	6.9 ± 0.4	0
Lys41	13.6 ± 0.7	15 ± 1	0

This table is continued on the next page.

Lys58	16.1 ± 0.7	16 ± 3	0
Lys75	1.6 ± 0.2	1.3 ± 0.1	0
Asn83	1.50 ± 0.04	1.4 ± 0.1	0
Lys91	27 ± 1	22 ± 1	- 18
Lys94	3.1 ± 0.2	3.4 ± 0.3	0
DEPC			
N-terminus	99 ± 1	68 ± 1	- 43
Thr4	86.7 ± 0.7	87 ± 2	0
Lys6	8.07 ± 0.05	8.0 ± 0.1	0
His13	47 ± 2	46 ± 3	0
Lys19	22 ± 1	19 ± 2	0
Ty26	0.75 ± 0.07	0.79 ± 0.05	0
Ser28	0.24 ± 0.01	0.17 ± 0.02	- 29
His31	1.8 ± 0.1	0.7 ± 0.1	- 61
Ser33	1.7 ± 0.1	1.1 ± 0.1	- 35
Lys41	0.32 ± 0.04	0.33 ± 0.03	0
His51	61 ± 3	61 ± 4	0
Ser57/Lys58	43 ± 2	42 ± 3	0
Tyr63	7.1 ± 0.5	4.3 ± 0.3	- 39
Tyr67	2.1 ± 0.1	1.75 ± 0.06	- 17
Lys75	0.36 ± 0.04	0.33 ± 0.02	0
Ser88	63 ± 3	65 ± 3	0
Lys94	30 ± 2	29 ± 2	0
BD			
Arg3	0.6 ± 0.1	1.05 ± 0.05	+ 75
Arg12	1.3 ± 0.1	1.00 ± 0.05	- 23
Arg45	3.8 ± 0.3	3.6 ± 0.1	0
Arg97	2.8 ± 0.4	4.8 ± 0.2	+ 71

3.3.4 Copper Binding to Monomeric β2m

Studies have shown that *in vitro* interactions of β2m with stoichiometric amounts of Cu(II) under near physiological conditions result in fibril formation [6,15]. The exact structural changes undergone by the protein as a result of Cu(II) binding, however, are unclear. To gather some insight into possible structural changes associated with Cu(II) binding, the reactions of NHSA, DEPC, and BD with β2m in the presence and absence of Cu(II) were performed. Several factors affect the covalent labeling efficiency of amino acid side chains, but solvent accessibility is the most important [20]. Consequently, covalent labeling is especially effective at identifying residues involved in ligand binding because the solvent accessibilities of such residues typically undergo noticeable changes. Thus, by monitoring the modification extents of different residues within 2 min after addition of Cu(II) we can obtain some insight into the residues that are involved in binding Cu(II) as well as those that undergo changes in solvent accessibility.

The β2m residues that react with NHSA, DEPC, and BD are widely distributed along the polypeptide chain and on the surface of the protein. Thus, we cover most regions of the protein. The structural data obtained covers 24 of the protein's 99 residues and ~30% of the surface-exposed residues (Figure 3.5). For the three covalent labels, linear relationships are observed at the concentrations used. At these reagent concentrations the protein's global structure is maintained for both native β2m (Figures 3.1-3.3) and β2m in the presence of Cu(II) (Figure 3.6). The extent of β2m modification decreased upon addition of Cu(II) for all the different NHSA and DEPC concentrations used. This observation is consistent with previous studies of other proteins that show that Cu(II) binding protects residues from DEPC modification [16,17].

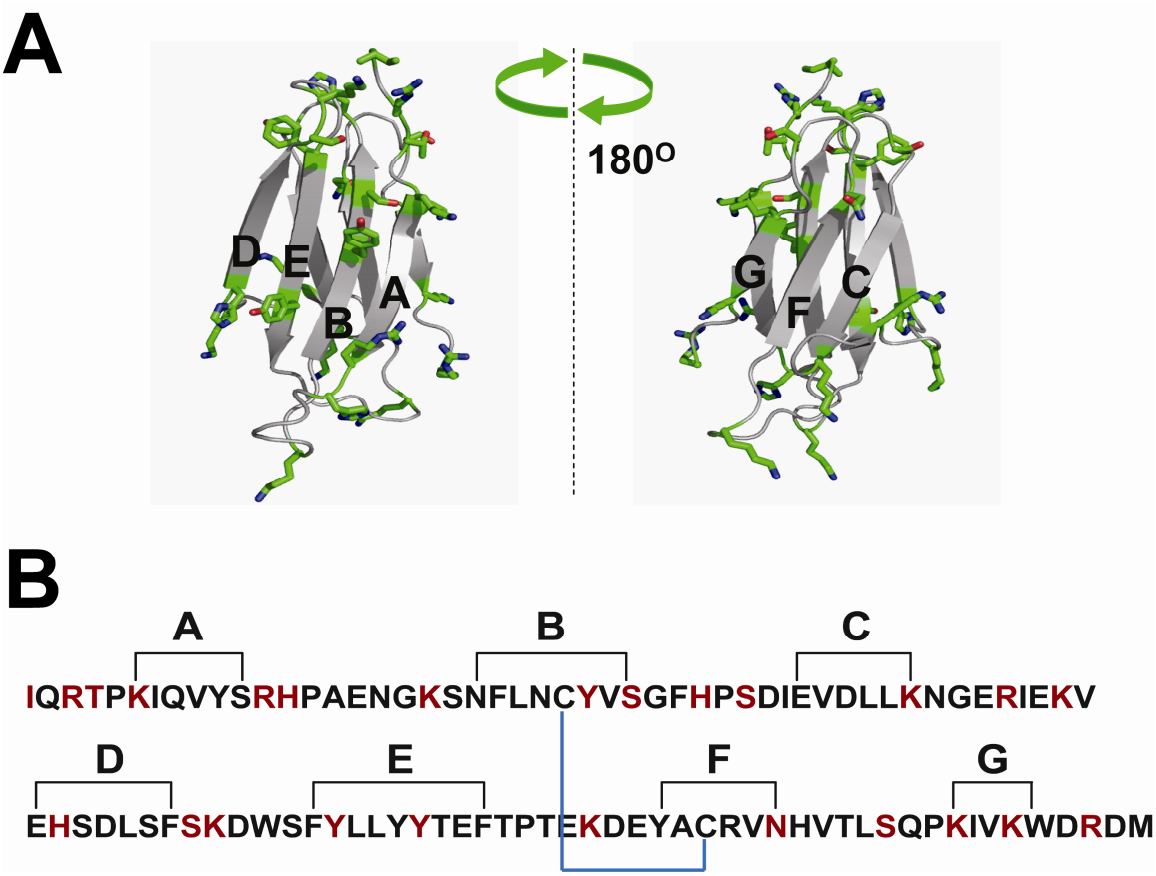


Figure 3.5 (A) Ribbon representation of monomeric β2m (PDB ID: 2D4F). Amino acids modified by the covalent labels are shown as sticks. (B) Amino acid sequence of β2m showing strand nomenclature [23]. Black lines show amino acids on each β strand. The internal disulfide bond is shown in blue. The amino acids probed by the covalent labels are shown in red.

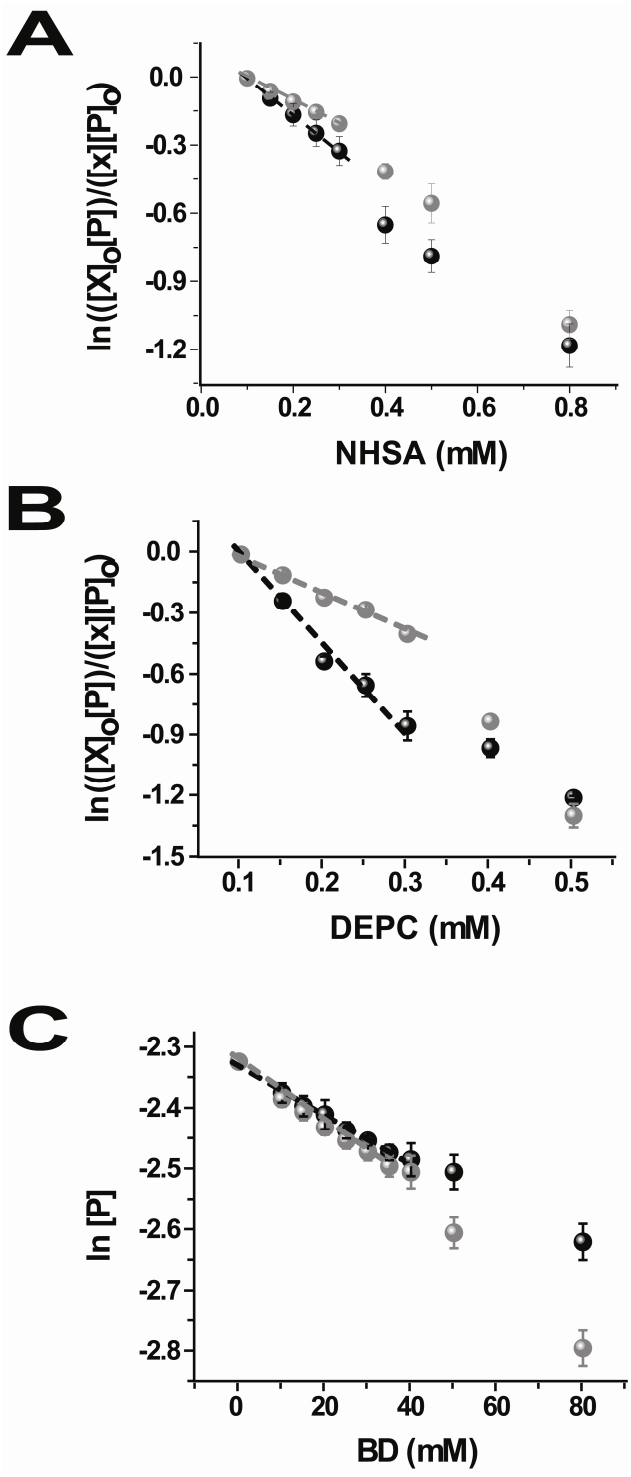


Figure 3.6 Dose-response plots for β2m after reaction with (A) DEPC, (B) NHSA, and (C) BD. The plots show that the concentration range in which the protein's global structure is maintained is the same for both native β2m and β2m in the presence of Cu(II).

Out of the 24 modified residues, only 10 amino acids are found to change in reactivity, namely, the N-terminus, Arg3, Arg12, Ser28, His31, Ser33, Tyr63, Tyr67, Lys91, and Arg97. This suggests that there are significant structural changes in the immediate regions surrounding these amino acids. These observations are consistent with previous studies that indicate that the N-terminus and His31 are part of the Cu(II) binding site in β2m [6,24]. The N-terminus and His31 showed the greatest decrease in reactivity, indicating that these residues are protected by Cu(II) from modification (Figure 3.7).

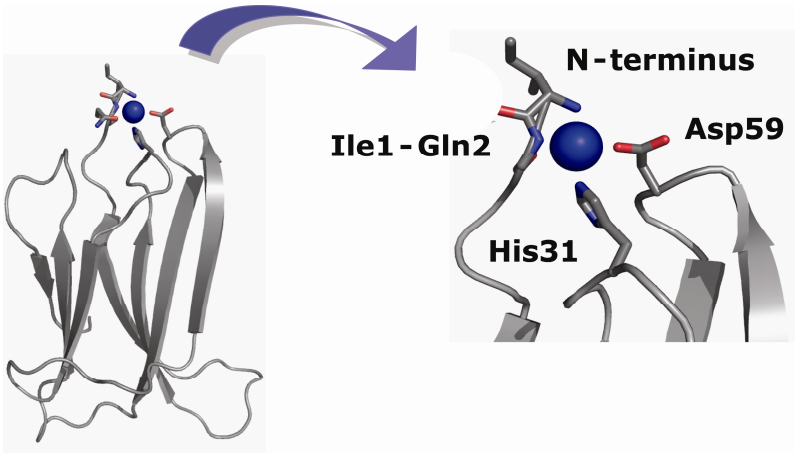


Figure 3.7 Structure of β2m showing the Cu(II) binding site at the N-terminal region of the protein. This structure was generated from molecular mechanics calculations of the protein, beginning with the crystal structure of the apo-protein (PDB 1LDS) [25] and after adding Cu(II) to the known binding sites.

Figure 3.8 shows that most of the remaining eight residues that change in reactivity are localized at the N-terminal region and the ABED β-sheet of β2m. The reduced reactivity of Ser33 might simply be explained by the nearby binding of Cu(II) to His31, which prevents ready access of DEPC to this amino acidic side chain. Alternatively, others have reported that Cu(II) binding induces a *cis* to *trans* backbone

isomerization of the Pro32 residue [26]. A β2m mutant (P32A) that closely resembles this isomerized structure was designed and its atomic structure was solved using X-ray diffraction (PDB 2F8O). In this crystal structure, the protein adopts a slightly different conformation that places Ser33 closer to the suspected Cu(II) binding site, which would make it less accessible and therefore less reactive. Thus, the drop in Ser33's reactivity observed in our studies may be caused by a similar structural change upon binding Cu(II).

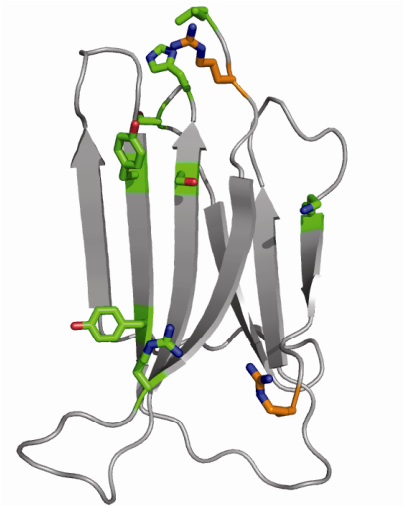


Figure 3.8 Ribbon representation of monomeric β2m (PDB ID 2D4F) [27]. Amino acids that decrease (green) and increase (orange) in modification extent upon interaction with Cu(II) are shown as sticks.

Our data also suggests that Cu(II) binding to monomeric β2m repositions Arg3. The increased reactivity of Arg3 might be explained by the repulsion of the positively charged side chain by Cu(II). Recently, a crystal structure of the H13F mutant of β2m in the presence of Cu(II) was solved, and it was found to form a hexamer with Cu(II) bound

at the N-terminal region of all β2m subunits [28]. Examination of one subunit of the H13F hexamer shows that the guanidinium group of Arg3 is pointing away from Cu(II) by approximately 11 Å (Figure 3.9). In the crystal structure of wild-type protein without Cu(II), Arg3 is located very close to His31 and where Cu(II) ultimately binds in the monomer. The location of Arg3 in the H13F hexamer suggests that Cu(II) binding forces this residue to be repositioned. Moreover, Arg3 is now located about 3.8 Å from Glu16 of another protein subunit, suggesting that these residues probably form an intermolecular salt bridge. Interestingly, this is not the only intermolecular salt bridge created by Cu(II) binding. A recent study using metal-catalyzed oxidation (MCO) revealed that Asp59 is part of the Cu(II) binding site. In both NMR and crystal structures of monomeric β2m [25,29], the side chain of Asp59 points away from the N-terminal strand. However, upon binding Cu(II) this residue must be repositioned to point toward the N-terminal amine and the amide between Ile1 and Gln2 to effectively bind Cu(II). A close examination of the crystal structure of the H13F hexamer of β2m [28] reveals that the side chain of Asp59 has rotated greater than 100º and is located closer to the N-terminus than in the monomer. Furthermore, inspection of the subunit contacts between chains B and C in the crystal structure of the H13F hexamer of β2m reveals that Asp59 of one protein subunit is about 2.8 Å from Lys19 of another protein subunit. The interactions of Arg3/Glu16 and Asp59/Lys19 are symmetric, thereby giving rise to a pair of salt bridges that are probably important for the stability of the β2m dimer. The formation of these salt bridges is highly probable because electrostatic interactions are usually important contributors to the specificity and stability of protein-protein complexes [30-32].

Figure 3.9 A comparison of the crystal structure of monomeric β2m (green, PDB 2D4F) and one of the subunits from the H13F hexamer (blue, PDB 3CIQ) that illustrates the repositioning of Arg3.

Lastly, four residues on the ABED β-sheet of β2m decreased in reactivity upon Cu(II) binding, namely, Arg12, Ser28, Tyr63, and Tyr67. In addition to the possible intermolecular salt bridges formed by Arg3/Glu16 and Asp59/Lys19, the rearrangements of these four residues on the ABED β-sheet likely help facilitate dimer formation of β2m. In a previous work, we have shown that a better indicator of reactivity appears to be the solvent accessibility of the reactive atom (i.e. hydroxyl group for Ser and Tyr and guanidinium group for Arg) [21]. The decrease in reactivity of these residues suggest that the side chains move after Cu(II) binding such that their reactive atoms become less solvent accessible. This movement probably exposes their hydrophobic part to the solvent, and together with the surrounding nonpolar residues, creates a hydrophobic facen on the ABED β-sheet. Studies of other amyloidogenic proteins showed that hydrophobic interactions play a decisive role in the formation of amyloid fibrils [33-35]. Hence, we hypothesize that this hydrophobic region can promote dimerization by interaction with

another ABED β-sheet to minimize solvent exposed hydrophobic groups. Examination of the subunit contacts between chains B and C in the H13F hexamer shows that these residues have indeed moved > 2Å relative to the wild-type apo β2m (PDB ID 2D4F) and that they formed hydrophobic interactions with the ABED β-sheet of an adjacent monomer.

The decrease and increase in reactivity of Lys91 and Arg97, respectively, are difficult to explain. The increase in modification of Arg97 indicates that it becomes more solvent accessible and that the C-terminal strand possibly becomes frayed upon Cu(II) binding. This is consistent with a study wherein β2m was shown to populate partially unfolded states under acidic conditions that display significant destabilization of the C-terminal strand [36] with the fraying of the C-terminal strand.

The fact that the reactivity of His51 is the same in the absence and presence of Cu(II) is also informative. The literature is contradictory with regards to His51's interaction with Cu(II). Our group and others [6,24] have found no evidence for Cu(II) binding to His51, whereas an NMR study [37] implies that Cu(II) does bind His51 based on the broadening of resonances associated with this residue. The present work provides additional evidence that Cu(II) does not interact with His51, perhaps indicating that the previous NMR results could simply be due to increased conformational flexibility at this residue when Cu(II) binds to β2m.

3.4 Conclusions

In conclusion, this study shows that the combination of selective covalent modification with mass spectrometric detection is well suited for the characterization metal binding and binding-induced conformational changes. Initial application of this method to study Cu(II)-β2m interactions provides additional insight into the effect of Cu(II) on this protein's structure. The NHSA and DEPC reactivity of β2m in the presence of Cu(II) confirms the involvement of the N-terminus and His31 in the binding site and further rules out the interaction of His51 with Cu(II).

The results described here also give us additional insight into how Cu(II) facilitates the β2m structural changes that lead to oligomer formation. In the monomer, Cu(II) binds to the N-terminus and His31. Binding to these residues causes several structural changes. We propose that movement of Arg3 allows a dimer-stabilizing salt bridge to be formed with Glu16 of an adjacent β2m. The movement of Arg12, Ser28, Tyr63, and Tyr67 on the ABED β-sheet possibly creates a hydrophobic face that can promote dimerization via interaction with another hydrophobic sheet of an adjacent monomer. Rearrangement of these residues, along with the *cis-trans* isomerization of Pro32, may therefore be part of the switch that enables Cu(II) to initiate dimer formation [26,38].

Lastly, the results in this chapter provide some additional details about the molecular role that Cu(II) can play in amyloid-forming systems. Divalent transition metals, especially Cu(II), appear to be one of the general ways to cause protein amyloid formation. Prior work shows that proteins such as α-synuclein, immunoglobulin light chain, Aβ, and even possibly the prion protein form amyloid aggregates in the presence

of metal ions. In the present work we find that metal binding to certain residues and repulsion of other residues can perturb protein structure in a way that enables aggregation. Further studies of Cu(II)-β2m interactions and the resulting aggregation may provide even more insight into the possible roles that divalent metals play in protein aggregation.

3.5 References

1. McParland, V. J., Kad, N. M., Kalverda, A. P., Brown, A., Kirwin-Jones, P., Hunter, M. G., Sunde, M., and Radford, S. E. (2000) Partially unfolded states of β_2-microglobulin and amyloid formation *in vitro*. *Biochemistry 39*, 8735-8746.

2. Esposito, G., Michelutti, R., Verdone, G., Viglino, P., Hernandez, H., Robinson, C. V., Amoresano, A., Dal Piaz, F., Monti, M., Pucci, P., Mangione, P., Stoppini, M., Merlini, G., Ferri, G., and Bellotti, V. (2000) Removal of the N-terminal hexapeptide from human β2-microglobulin facilitates protein aggregation and fibril formation. *Protein Sci. 9*, 831-845.

3. Relini, A., Canale, C., De Stefano, S., Rolandi, R., Giorgetti, S., Stoppini, M., Rossi, A., Fogolari, F., Corazza, A., Esposito, G., Gliozzi, A., and Bellotti, V. (2006) Collagen plays an active role in the aggregation of β_2-microglobulin under physiopathological conditions of dialysis-related amyloidosis. *J. Biol. Chem. 281*, 16521-16529.

4. Ohhashi, Y., Kihara, M., Naiki, H., and Goto, Y. (2005) Ultrasonication-induced amyloid fibril formation of β_2-microglobulin. *J. Biol. Chem. 280*, 32843-32848.

5. Morgan, C. J., Gelfand, M., Atreya, C., and Miranker, A. D. (2001) Kidney dialysis-associated amyloidosis: a molecular role for Cu(II) in fiber formation. *J. Mol. Biol. 309*, 339-345.

6. Eakin, C. M., Knight, J. D., Morgan, C. J., Gelfand, M. A., and Miranker, A. D. (2002) Formation of a copper specific binding site in non-native states of β-2-microglobulin. *Biochemistry 41*, 10646-10656.

7. Bush, A. I. and Tanzi, R. E. (2002) The galvanization of beta-amyloid in Alzheimer's disease. *Proc. Natl. Acad. Sci. USA 99*, 7317-7319.

8. Uversky, V. N., Li, J., and Fink, A. L. (2001) Metal-triggered structural transformations, aggregation, and fibrillation of human alpha-synuclein. A possible molecular link between Parkinson's disease and heavy metal exposure. *J. Biol. Chem. 276*, 44284-44296.

9. Wadsworth, J. D., Hill, A. F., Joiner, S., Jackson, G. S., Clarke, A. R., and Collinge, J. (1999) Strain-specific prion-protein conformation determined by metal ions. *Nature Cell Biol. 1*, 55-59.

10. Jobling, M. F., Huang, X., Stewart, L. R., Barnham, K. J., Curtain, C., Volitakis, I., Perugini, M., White, A. R., Cherny, R. A., Masters, C. L., Barrow, C. J., Collins, S. J., Bush, A. I., and Cappai, R. (2001) Copper and zinc binding modulates the aggregation and neurotoxic properties of the prion peptide PrP106-126. *Biochemistry 40*, 8073-8084.

11. Davis, D. P., Gallo, G., Vogen, S. M., Dul, J. L., Sciarretta, K. L., Kumar, A., Raffen, R., Stevens, F. J., and Argon, Y. (2001) Both the environment and somatic mutations govern the aggregation pathway of pathogenic immunoglobulin light chain. *J. Mol. Biol. 313*, 1021-1034.

12. Eakin, C. M. and Miranker, A. D. (2005) From chance to frequent encounters: origins of β2-microglobulin fibrillogenesis. *Biochim. Biophys. Acta 1753*, 92-99.

13. Antwi, K.; Mahar, M.; Srikanth, R.; Olbris, M. R.; Tyson, J. F.; Vachet, R. W. (2008) Cu(II) organizes β-2-microglobulin oligomers but is released upon amyloid formation. *Protein Sci. 17*, 748-759.

14. Calabrese, M. F. and Miranker, A. D. (2007) Formation of a stable oligomer of β-2 microglobulin requires only transient encounter with Cu(II). *J. Mol. Biol. 367*, 1-7.

15. Eakin, C. M., Attenello, F. J., Morgan, C. J., and Miranker, A. D. (2004) Oligomeric assembly of native-like precursors precedes amyloid formation by β-2 microglobulin. *Biochem*istry *43*, 7808-7815.

16. Qin, K. F., Yang, Y., Mastrangelo, P. and Westaway, D. (2002) Mapping Cu(II) binding sites in prion proteins by diethyl pyrocarbonate modification and matrix-assisted laser desorption ionization-time of flight (MALDI-TOF) mass spectrometric footprinting. *J. Biol. Chem. 277*, 1981-1990.

17. Narindrasorasak, S., Kulkarni, P., Deschamps, P., She, Y. M. and Sarkar, B. (2007) Characterization and copper binding properties of human COMMD1 (MURR1). *Biochemistry 46*, 3116-3128.

18. Fliss, H., and Viswanatha, T. (1979) 2,3-Butanedione as a photosensitizing agent – Application to alpha-amino acids and alpha-chymotrypsin. *Can. J. Biochem. 57*, 1267-1272.

19. Riordan, J. F. (1979) Arginyl residues and anion binding sites in proteins. *Mol. Cell Biochem. 26*, 71-92.

20. Mendoza, V. L., and Vachet, R. W. (2009) Probing protein structure by amino acid-specific covalent labeling and mass spectrometry. *Mass Spec. Rev. 28*, 785-815.

21. Mendoza, V. L., and Vachet, R. W. (2008) Protein surface mapping using diethylpyrocarbonate with mass spectrometric detection. *Anal. Chem. 80*, 2895-2904.

22. Hassani, O., Mansuelle, P., Cestele, S., Bourdeaux, M., Rochat, H., and Sampieri, F. (1999) Role of lysine and tryptophan residues in the biological activity of toxin VII (Tsγ) from the scorpion Tityus Serrulatus. *Eur. J. Biochem. 260*, 76-86.

23. Okon, M., Bray, P., and Vucelic, D. (1992) H^1-NMR assignments and secondary structure of human beta-2-microglobulin in solution. *Biochemistry 31*, 8906−8915.

24. Lim, J. and Vachet, R. W. (2004) Using mass spectrometry to study copper-protein binding under native and non-native conditions: β-2-microglobulin. *Anal. Chem. 76*, 3498-3504.

25. Trinh, C. H.; Smith, D. P.; Kalverda, A. P.; Phillips, S. E. V. and Radford, S. E. (2002) Crystal structure of monomeric human β-2-microglobulin reveals clues to its amyloidogenic properties. *Proc. Natl. Acad. Sci. USA 99*, 9771-9776.

26. Eakin, C. M., Berman, A. J., and Miranker, A. D. (2006) A native to amyloidogenic transition regulated by a backbone trigger. *Nature Struct. Mol. Biol. 13*, 202-208.

27. Kihara, M., Chatani, E., Iwata, K., Yamamoto, K., Matsuura, T., Nakagawa, A., Naiki, H., and Goto, Y. (2006) Conformation of amyloid fibrils of β2-microglobulin probed by tryptophan mutagenesis. *J. Biol. Chem. 281*, 31061-31069.

28. Calabrese, M. F., Eakin, C. M., Wang, J. M., and Miranker, A. D. (2008) A regulatable switch mediates self-association in an immunoglobulin fold. *Nature Struct. Mol. Biol. 15*, 965-971.

29. Verdone, G.; Corazza, A.; Viglino, P.; Pettirossi, F.; Giorgetti, S.; Mangione, P.; Andreola, A.; Stoppini, M.; Bellotti, V. and Esposito, G. (2002) The solution structure of human β2-microglobulin reveals the prodromes of its amyloid transition. *Protein Sci. 11*, 487-499.

30. Xu, D.; Lin, S. L. and Nussinov, R. (1997) Protein binding *versus* protein folding: The role of hydrophilic bridges in protein associations. *J. Mol. Biol. 265*, 68-84.

31. Sheinerman, F. B.; Norel, R. and Honig, B. (2000) Electrostatic aspects of protein-protein interactions. *Curr. Opin. Struct. Biol. 10*, 153-159.

32. Keskin, O.; Gursoy, A.; Ma, B. and Nussinov, R. (2008) Principles of protein-protein interactions: What are the preferred ways for proteins to interact? *Chem. Rev. 108*, 1225-1244.

33. Kim, W., and Hecht, M.H. (2006) Generic hydrophobic residues are sufficient to promote aggregation of the Alzheimer's A beta 42 peptide. *Proc. Natl. Acad. Sci. USA 103*, 15824-15829.

34. Kammerer, R.A., Kostrewa, D., Zurdo, J., Detken, A., Garcia-Echeverria, C., Green, J.D., Muller, S.A., Meier, B.H., Winkler, F.K., Dobson, C.M., and Steinmetz, M.O. (2004) Exploring amyloid formation by a de novo design. *Proc. Natl. Acad. Sci. USA 101*, 4435-4440.

35. Giasson, B.I., Murray, I.V.J., Trojanowski, J.Q., and Lee, V.M.Y. (2001) A hydrophobic stretch of 12 amino acid residues in the middle of alpha-synuclein is essential for filament assembly. *J. Biol. Chem. 276*, 2380-2386.

36. McParland, V.J., Kalverda, A.P., Homans, S.W., and Radford, S.E. (2002) Structural properties of an amyloid precursor of β_2-microglobulin. *Nature Struct. Biol. 9*, 326-331.

37. Villanueva, J., Hoshino, M., Katou, H., Kardos, J., Hasegawa, K., Naiki, H., and Goto, Y. (2004) Increase in the conformational flexibility of β_2-microglobulin upon copper binding: a possible role for copper in dialysis-related amyloidosis. *Protein Sci. 13*, 797-809.

38. Eichner, T., and Radford, S.E. (2009) A generic mechanism of β2-microglobulin amyloid assembly at neutral pH involving a specific proline switch. *J. Mol. Biol. 386*, 1312-1326.

CHAPTER 4

STRUCTURE OF THE PRE-AMYLOID DIMER OF β-2-MICROGLOBULIN FROM COVALENT LABELING AND MASS SPECTROMETRY

This chapter is part of a paper published as: Mendoza, V.L., Antwi, K., Baron-Rodriguez, M.A., Blanco, C., and Vachet, R.W. (2009) Structure of the pre-amyloid dimer of β-2-microglobulin from covalent labeling and mass spectrometry. *Biochemistry 49*, 1522-1532.

4.1 Introduction

The Cu(II)-induced amyloid formation of β-2-microglobulin (β2m) is preceded by the formation of discrete, oligomeric intermediates [1,3]. The formation of discrete oligomers as prefibrillar intermediates has been observed for several other amyloidogenic proteins too [4-6], and indeed acid-induced (pH = 2.5) amyloid formation of β2m has been found to be preceded by discrete oligomers [7]. Delineating the mechanism of the early stages of β2m fibrillogenesis and the structural properties of these oligomeric intermediates may be critical for developing therapeutic strategies against DRA. Furthermore, recent studies of other amyloid systems suggest that prefibrillar intermediates might be responsible for cellular toxicity rather than the amyloid fibrils [8,9]. Consequently, any methods or results that give insight into intermediates that precede β2m fibrils may be useful for understanding other amyloid systems as well.

In this chapter, we used Cu(II) as a trigger to initiate β2m amyloid formation, and then studied the structure of the dimer, which is the first pre-amyloid oligomer formed [1,3]. Previous work has shown that incubation of β2m with Cu(II) under near physiological conditions results in the formation of dimers in the first 30 minutes with tetramers and hexamers not forming for > 12-24 hours [1,3]. Because the dimer is an

intermediate and is always present as a mixture with the monomer and eventually other higher order oligomers, obtaining residue-specific information about the dimer is very challenging by traditional approaches. One recently used approach to solve this problem has been to construct β2m mutants that are capable of crystallizing as oligomers [10,11]. In this way, it has been found that the P32A mutant of β2m forms a dimer in the absence of Cu and the H13F mutant forms a stable hexamer in the presence of Cu. These crystal structures have then been used to support hypotheses about the structural changes necessary for β2m oligomer and amyloid formation. While these crystal structures provide high resolution atomic-level information, they cannot provide a complete picture of β2m oligomer structure. The β2m mutants do not form fibrils, and so the crystallized oligomers unlike the wild-type oligomers contain structural features that prevent further aggregation. Moreover, it is possible that the selected mutations initiate the formation of oligomers that differ completely from the ones formed by the wild-type protein.

As a complementary approach to these crystallographic studies, we describe here the use of covalent labeling along with mass spectrometry (MS) to gather insight into the structure of the wild-type β2m dimer. This method relies on amino acid specific reactions of solvent exposed amino acid side chains, and tandem MS (MS/MS) to identify the modified residues. Covalent labeling combined with MS and MS/MS has been used to map protein surfaces, identify ligand-binding sites, study protein-protein and protein-nucleic acid complexes, and detect ligand-induced conformational changes [12]. These covalent labeling methods provide information about protein structure and interactions by comparing the reactivity of amino acid side chains in solution under two or more different conditions. MS/MS is then used to identify the specific amino acids that

undergo significant changes in reactivity. This approach is well suited for providing insight into the dimeric intermediate of β2m because amino acid side chains are usually buried upon formation of new protein-protein interfaces, and buried amino acids are much less reactive than exposed amino acids. Furthermore, the specificity of MS allows this information to be obtained even under conditions in which a mixture of β2m forms is present.

4.2 Experimental Procedure

4.2.1 Materials

The same materials were used as described in Section 3.2.1.

4.2.2 Formation of β2m Oligomers

Reports have shown that discrete oligomers precede β2m amyloid fibrils when monomeric β2m is incubated under near-physiological conditions in the presence of Cu(II) [1,3]. Amyloid fibrils were formed by incubation of 100 μM β2m in 200 mM potassium acetate, 500 mM urea, and 25 mM MOPS (pH 7.4) with 200 μM copper(II) sulfate at 37 °C. All components were equilibrated at 37 °C prior to Cu(II) addition, and immediately returned to 37 °C after mixing. The covalent labels were added to aliquots of incubated β2m taken at several time points after initiating the amyloid fibril formation reaction.

4.2.3 Covalent Labeling Reactions

The covalent labeling reactions were done in a manner identical to that described in Sections 3.2.3-3.2.5.

4.2.4 Proteolytic Digestion

The proteolytic digestions were performed in a manner identical to that described in Section 3.2.6.

4.2.5 Desalting

Desalting of the protein samples just prior to electrospray ionization (ESI) mass spectrometry (MS) analyses was performed using a 5 ml HiTrap desalting column from Amersham Biosciences (Piscataway, NJ). The column was first equilibrated with 30 ml of the mobile phase (10 mM ammonium acetate, pH = 7.4) at a flow rate of 5 mL/min. Then, between 100 and 150 μL of incubated sample was injected onto the column via a sample loop attached to a home-made setup. The oligomeric intermediates were eluted at a flow rate of 10 ml/min using 10 mM ammonium acetate (pH = 7.4), which was filled into a syringe and delivered by a syringe pump. The β2m fraction eluting from the column was collected between 20 and 22 seconds and analyzed immediately by ESI-MS.

4.2.6 Instrumentation

The amount of modification was determined by removing an aliquot of the purified β2m and analyzing the samples using a Bruker Esquire-LC quadrupole ion trap mass spectrometer (Billerica, MA) equipped with an ESI source. The ESI source was

operated at a spray voltage of 3.5 kV, and the capillary temperature was set at 300°C. The voltages for the transfer optics between the ESI source and the ion trap were optimized for maximum signal, with typical skimmer 1 and capillary offset values of 30-35 V and 50-60 V, respectively.

The proteolytic fragments were separated by an Agilent HP1100 (Wilmington, DE) HPLC system with a C18 column (15 cm x 2.1 mm, 5 μm particle size, Supelco, St. Louis, MO) for on-line analysis by MS and MS/MS. The fragments of NHSA- and BD-modified β2m were eluted using a linear gradient of methanol that increased from 5 to 70% over 20 min and 70 to 100% over the final 2 min at a flow rate of 0.250 mL/min. For the fragments of DEPC-modified β2m, a linear gradient of methanol that increased from 5 to 70% over 30 min and 70 to 100% over the final 3 min was used. For both gradient conditions, water comprised the balance of the solvent, and a total of 0.1% acetic acid was present. The LC effluent was fed into the mass spectrometer with similar ESI source conditions as described above. Tandem mass spectra were acquired using collision induced dissociation (CID) with isolation widths of 1.0 Da and excitation voltages between 0.6 and 1.0 V. Peptide sequences were determined from the MS/MS data via *de novo* sequencing or with the help of BioTools (Bruker Daltonics, Billerica, MA).

Sequencing of β2m to identify the specific amino acids modified by BD is not possible using collision induced dissociation (CID), so electron transfer dissociation (ETD) was used to identify the specific modification sites. In the CID process the BD label is readily lost without any associated information about its location in the sequence. The ETD experiments were carried out using a Bruker HCT ETD II (Billerica, MA) equipped with an ESI source. The voltages for the transfer optics between the ESI source

and the ion trap were optimized for maximum signal. ETD was performed using reaction time of 25 ms and low mass cutoff of 100. Accumulation times for the fluoranthene radical anion, which was used as the ETD reagent, were typically between 3 and 5 ms.

A JMS-700 MStation (JEOL, Tokyo, Japan) double-focusing mass spectrometer equipped with a standard ESI source was used to acquire mass spectra of the β2m oligomers after desalting the incubation solution. The ESI source conditions that were used are the following: desolvating plate temperature, 100 ^{o}C; orifice temperature, 140 ^{o}C; orifice potential, 60 V; ring lens potential, 100 V. ESI-MS experiments of the desalted protein samples were used to measure the average charge states of these oligomers and to estimate their surface areas using a method described previously by Kaltashov and co-workers [13]. The spectra were acquired by scanning the magnet at a rate of 5 s/decade.

4.2.7 Amino Acid Modification Percentage

The amino acid modification percentages were calculated in a manner identical to that described in Section 3.2.8.

4.2.8 Determination of Solvent Accessibility

The 3D structures of β2m that were examined were obtained from the Protein Data Bank: 2D4F for monomeric β2m [14], 2F8O for the P32A dimer [10], and 3CIQ for the H13F hexamer [11]. The solvent accessible surface areas (SASA) of individual amino acid side chains were calculated from the PDB coordinate files using GETAREA (http://curie.utmb.edu/getarea.html) [15]. For these SASA calculations, a 4.0 Å probe

radius was used because the radii of the covalent labels used in this study range from ~3.5-4.5 Å. To determine the surface area of the whole protein for comparison with the ESI-MS results, a 1.4 Å probe radius was used.

4.2.9 Docking Calculations

All simulations were performed using the β2m subunit taken from the crystal structure of human MHC-I complex (PDB ID: 1DUZ) [16], which contains 99 amino acid residues. Met0 was eliminated from the original set of coordinates in order to resemble the wild-type protein used in the labeling experiments. Hydrogen atoms were explicitly added to the protein model and His residues were considered to be neutral, as shown in nuclear magnetic resonance (NMR) studies by Esposito and co-workers [17]. Water molecules, as a TIP3P model [18], were also added until a cubic simulation cell of side length 7.8 nm was achieved.

Five ns molecular dynamics (MD) simulations were performed to determine the most probable conformation of wild-type β2m. The MD simulations were run at 310 K with periodic boundary conditions in an nPT ensemble and with the AMBER94 force field [19,20] implemented in GROMACS [21-23]. Protein structures were selected every 1 ps, and those structures that had root-mean-squared displacements of the backbone less than 0.05 nm were clustered together. The final structure that was used in the docking calculations was chosen from the largest cluster of structures.

The docking of two wild-type monomeric β2m molecules was performed to evaluate possible dimer interfaces. The docking search optimizes desolvation, grid-based shape complementarity, and electrostatics using a Fast Fourier Transform algorithm

ZDOCK2.3 [24,25] and provides a score for all possible docked conformations. The structures of the ten top-scoring dimers were then energy-minimized with solvent present using the AMBER94 force field. The Adaptive Poisson-Boltzmann Solver [26] within the AMBER force field was used to evaluate the free energy binding ($\Delta\Delta G$) of each dimer.

4.3 Results and Discussion

All covalent labeling experiments were performed on wild-type protein under solution conditions in which β2m forms amyloid fibrils. In previous studies, dynamic light scattering, size-exclusion chromatography, and MS showed that β2m oligomers are formed sequentially with the dimer first appearing between 30 min and 1 h and the initial form of the tetramer appearing after approximately 12-24 h [3]. These previous studies also showed that the smaller oligomers are in equilibrium with one another and can therefore dissociate to lower-order oligomers or to the monomer. Because complete isolation of the dimer was not possible without some dissociation back to the monomer, the chemical modification reactions were performed at different times before (t = 0 min) and after (t = 2 min, 0.5 h, 1 h, 1.5 h, and 2 h) adding Cu(II) to initiate the amyloid-forming reaction. The dimer percentage at time points after the addition of Cu(II) increases from 0 to 20% based on ESI-MS measurements of the desalted sample under amyloid-forming conditions, as described previously [*3*].

4.3.1 Covalent Labeling with NHSA

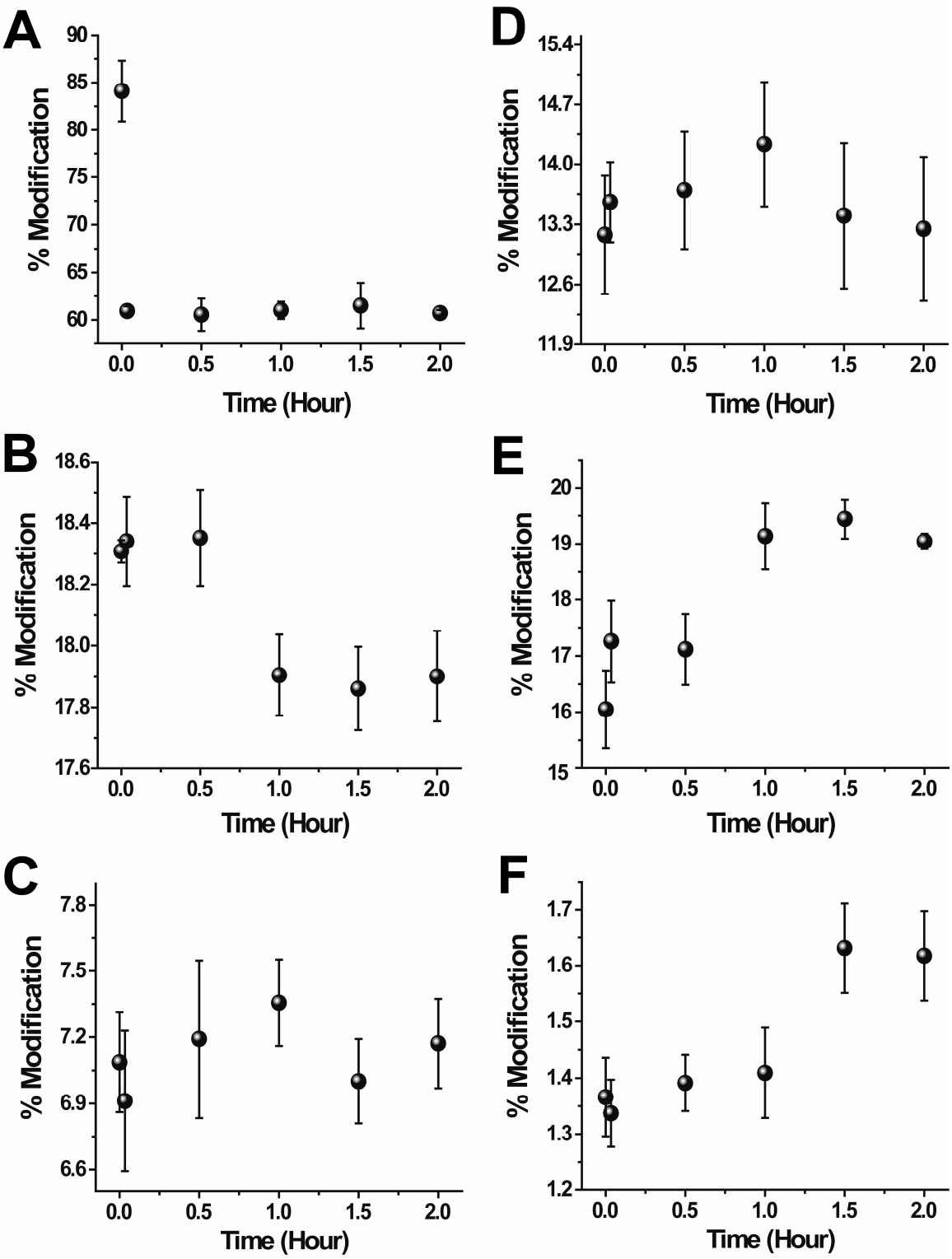
Amino groups such as the ε-NH_2 of lysine residues and the N-terminal α-NH_2 can react with NHSA. Identifying the amino acids that are acetylated is necessary to pinpoint the specific residues that become less reactive as the dimer concentration increases, and

this identification was achieved by proteolytic digestion and LC-MS/MS analyses. LC-MS/MS analyses indicate that the N-terminus, almost all of the lysines (Lys6, Lys19, Lys41, Lys58, Lys75, Lys91, and Lys94), and Asn83 are labeled to different degrees. The unmodified and modified fragments containing the lysine residues were detectable at all time points; however, the level of modification for some residues changed over time (*vide infra*). When the covalent labeling reactions with NHSA are performed at different time points after initiating the formation of the dimer, the modification extents of some residues change as the concentration of the dimer increases in solution (Figure 4.1). The data in these figures are obtained by measuring the modification percentage, as described in the experimental section, for the peptide fragments containing the indicated residues. For each time point, the modification reaction was repeated three times. Because the percent modification is determined using the ion intensity ratios of the unmodified and modified peptide fragments, small changes in modification levels can be accurately and precisely determined.

From the data in Figure 4.1, we find that the levels of modification of Lys58, Lys75, and Lys94 increase by about 20%, 20%, and 40%, respectively, from before Cu(II) is added to 2 h after Cu(II) is added. In contrast, the reactivity of the N-terminus, Lys6, and Lys91 decrease after addition of Cu(II). The N-terminus has the greatest drop in extent of modification, decreasing 30% immediately upon addition of Cu(II). The reactivity of this residue, however, does not change anymore as the dimer forms in solution. Like the N-terminus, the reactivity of Lys91 initially decreases, about 20%, after addition of Cu(II); however, there is a small 5% increase in the extent of modification from 2 min to 2 h as the dimer's concentration increases in solution. Lys6 behaves

differently than the N-terminus and Lys91 in that its reactivity does not drop immediately upon addition of Cu(II) but instead decreases as the dimer's concentration increases (Figure 4.1B). The reactivity of Lys19, Lys41, and Asn83 remain approximately constant up to 2 h after the addition of Cu(II).

In summary, the NHSA reactivity data indicate that of the nine modified residues in β2m four of them undergo a notable change in reactivity as the dimer is formed in solution. The modification extents of Lys58, Lys75, and Lys94 increase, whereas the reactivity of Lys6 decreases. The N-terminus and Lys91 also decrease in reactivity, but their lower reactivity appears to be mostly due to Cu binding rather than the formation of dimer.



This figure is continued on the next page.

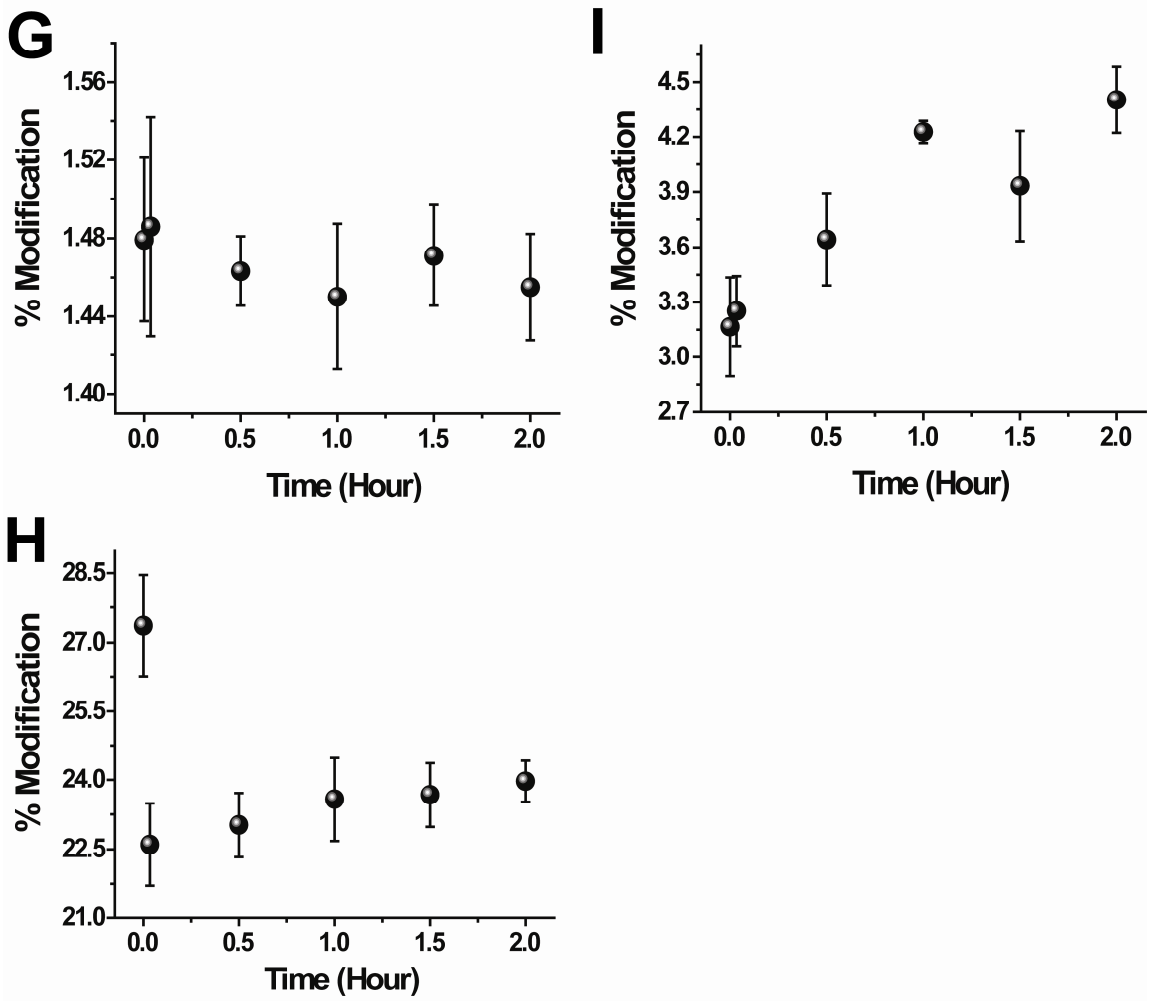


Figure 4.1 Extent of NHSA modification throughout the course of dimer formation. (A) N-terminus. (B) Lys6. (C) Lys19. (D) Lys41. (E) Lys58. (F) Lys75. (G) Ans83. (H) Lys91. (I) Lys94.

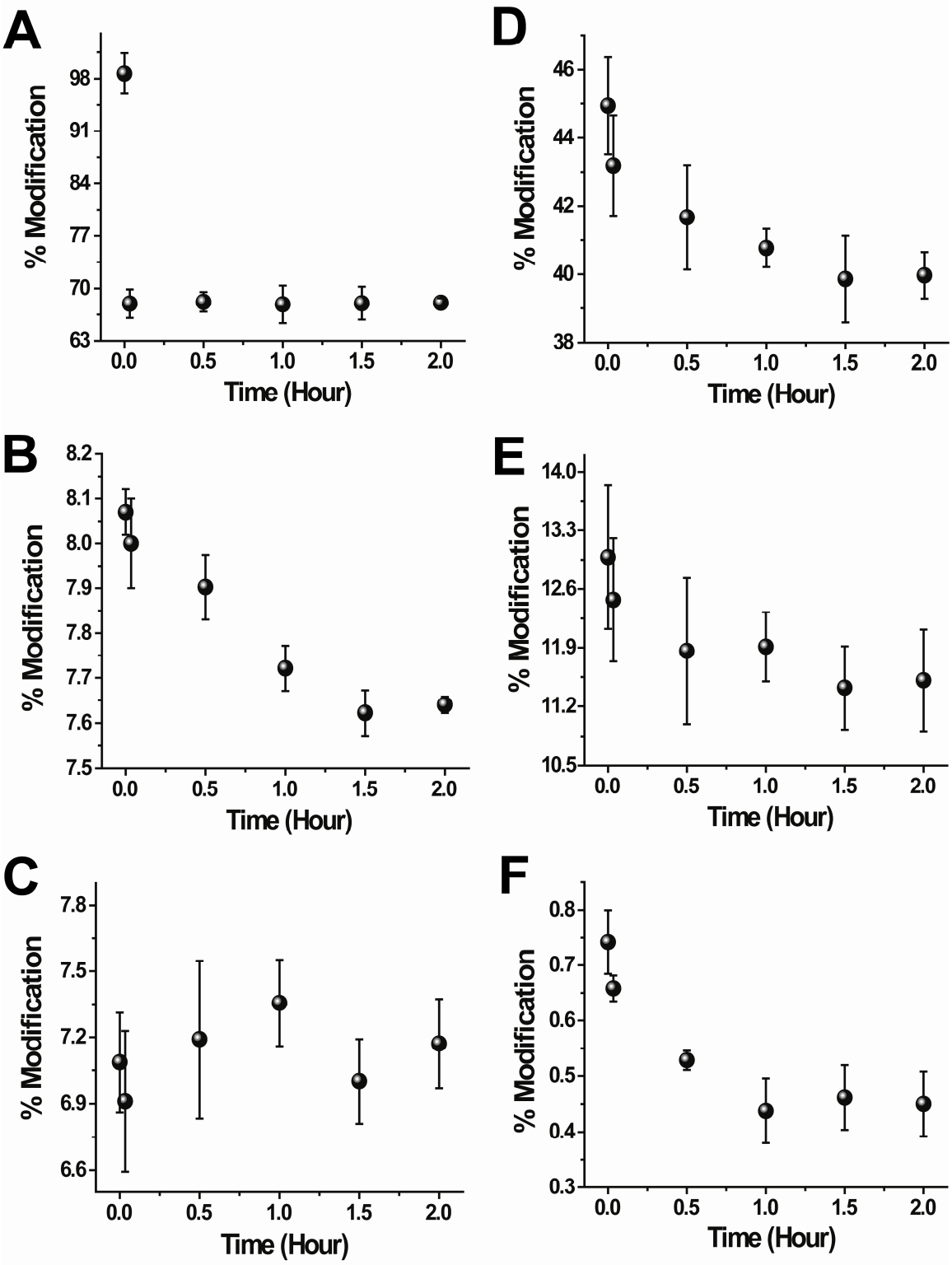
4.3.2 Covalent Labeling with DEPC

DEPC reacts readily with histidine residues but can also react with amine and hydroxyl groups at neutral pH. Proteolytic digestion of the protein and LC-MS/MS analyses indicate that DEPC reacted with the N-terminus, Thr4, His13, Lys19, Tyr26, Ser28, His31, Ser33, His51, Ser57/Lys58, Tyr63, Tyr67, Lys75, Ser88, and Lys94. These modifications were observed at all time points, but the level of modification for some

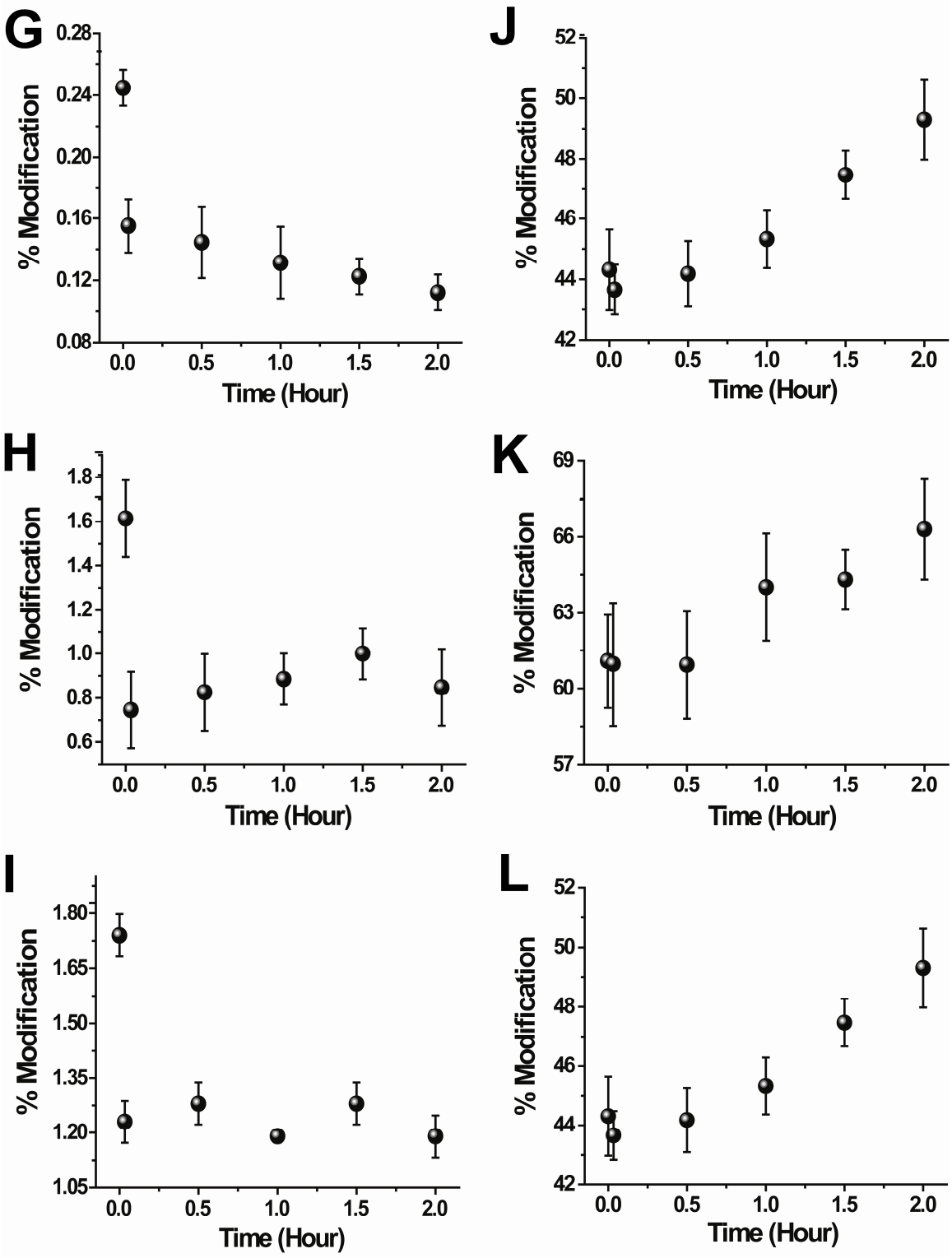
residues changed over time (*vide infra*). In all cases but one (Ser57/Lys58), the specific amino acids that were modified could be determined unambiguously from the MS/MS data.

The reactions with DEPC reveal that some residues change in reactivity as the dimer's concentration increases in solution (Figure 4.2). LC-MS data reveal that the percent modifications of six residues increase upon dimer formation. These residues are Thr4, His51, Ser57/Lys58, Lys75, and Lys94. Thr4 undergoes the least significant change in reactivity, with about a 3% increase when about 20% dimer is present. His51 and either Ser57 or Lys58 increase in reactivity by about 10%. The increased reactivity of Ser57 or Lys58 is consistent with the increased reactivity of Lys58 in the NHSA experiments described above. The reactivity of Lys75 and Lys94 increase the most, 25% and 30%, respectively, as the dimer's concentration increases. The significantly increased reactivities of Lys75 and Lys94 agree with the NHSA data described above. Finally, the reactivity trend of Ser88 is somewhat unclear. For all but the last time point (2 h), this residue reacts with DEPC in a fairly constant manner.

The modification extents of nine amino acids drop after Cu is added, but only five of these residues show decreased reactivity as the dimer's concentration increases. The reactivities of the N-terminus, His31, Ser33, and Tyr67 drop upon Cu(II) binding but change little or none as the dimer forms. Lys6, His13, Lys19, Tyr26, Ser28, and Tyr63, however, do decrease in reactivity as more and more dimer is formed.



This figure is continued on the next page.



This figure is continued on the next page.

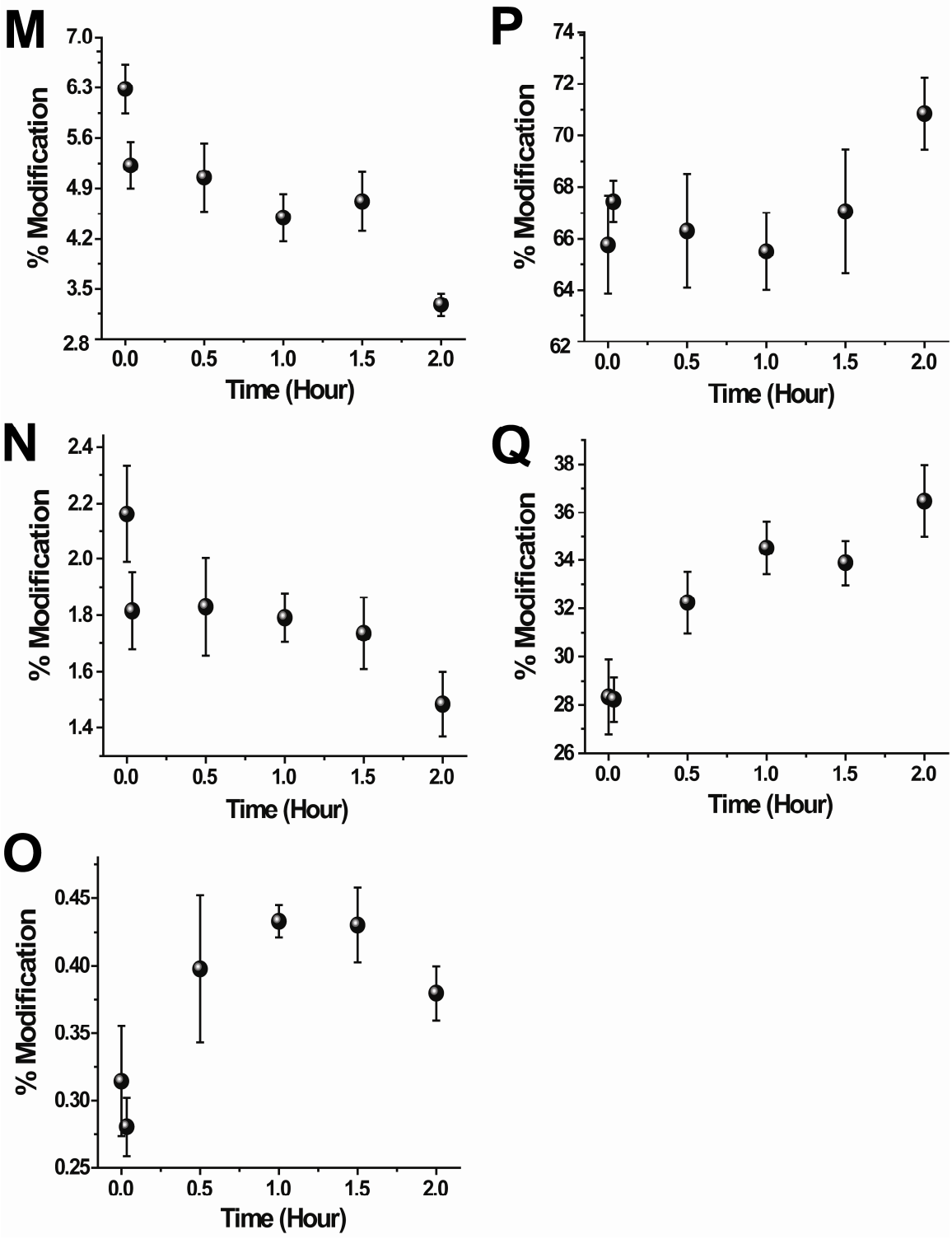


Figure 4.2 Extent of DEPC modification throughout the course of the dimer formation. (A) N-terminus. (B) Th4. (C) Lys6. (D) His13. (E) Lys19. (F) Tyr26. (G) Ser28. (H) His31. (I). Ser33. (J) Lys41. (K) His51. (L) Ser57/Lys58. (M) Tyr63. (N) Tyr67. (O) Lys75. (P) Ser88. (Q) Lys94.

The results from the DEPC reactions are consistent with the NHSA experiments, except for the reactivity trend of Lys19. In contrast to the constant modification of fragment Ser11-Phe22 with NHSA during the course of the dimer forming reaction, this fragment showed a 10% decrease in modification of Lys19 with DEPC over time (Figure 4.2E). This slight discrepancy might be due the microenvironment around Lys19, which could influence the reactivity of the two differently-charged labels to an unequal extent.

In summary, our results indicate that the reactivities of DEPC with His13, Lys19, Tyr26, Ser28, and Tyr63 decrease as the dimer is formed, whereas the reactivities of Thr4, His51, Ser57/Lys58, Lys75, and Lys94 increase as the dimer is formed. The modification extents for the N-terminus, His31, Ser33, and Tyr67 also decrease, but their lower reactivity appears to be due to Cu binding rather than the formation of dimer.

4.3.3 Covalent Labeling with Butanedione

Unlike DEPC, which reacts with many nucleophilic groups, BD reacts specifically with arginine. The LC/MS and ETD data indicate that BD reacts with Arg3, Arg12, Arg45, and Arg97 but not Arg81. The reactions of BD with β2m reveal that each of the four reacting arginine residues undergoes changes in reactivity after the addition of Cu(II) (Figure 4.3). The reactivities of Arg3 and Arg97 dramatically increase (about 80%) immediately after Cu(II) is added but do not change much as the dimer's concentration increases. The reactivity trends of Arg12 and Arg45, however, do change as the dimer is formed. The modification extent of Arg12 decreases over time, whereas the modification extent of Arg45 gradually increases.

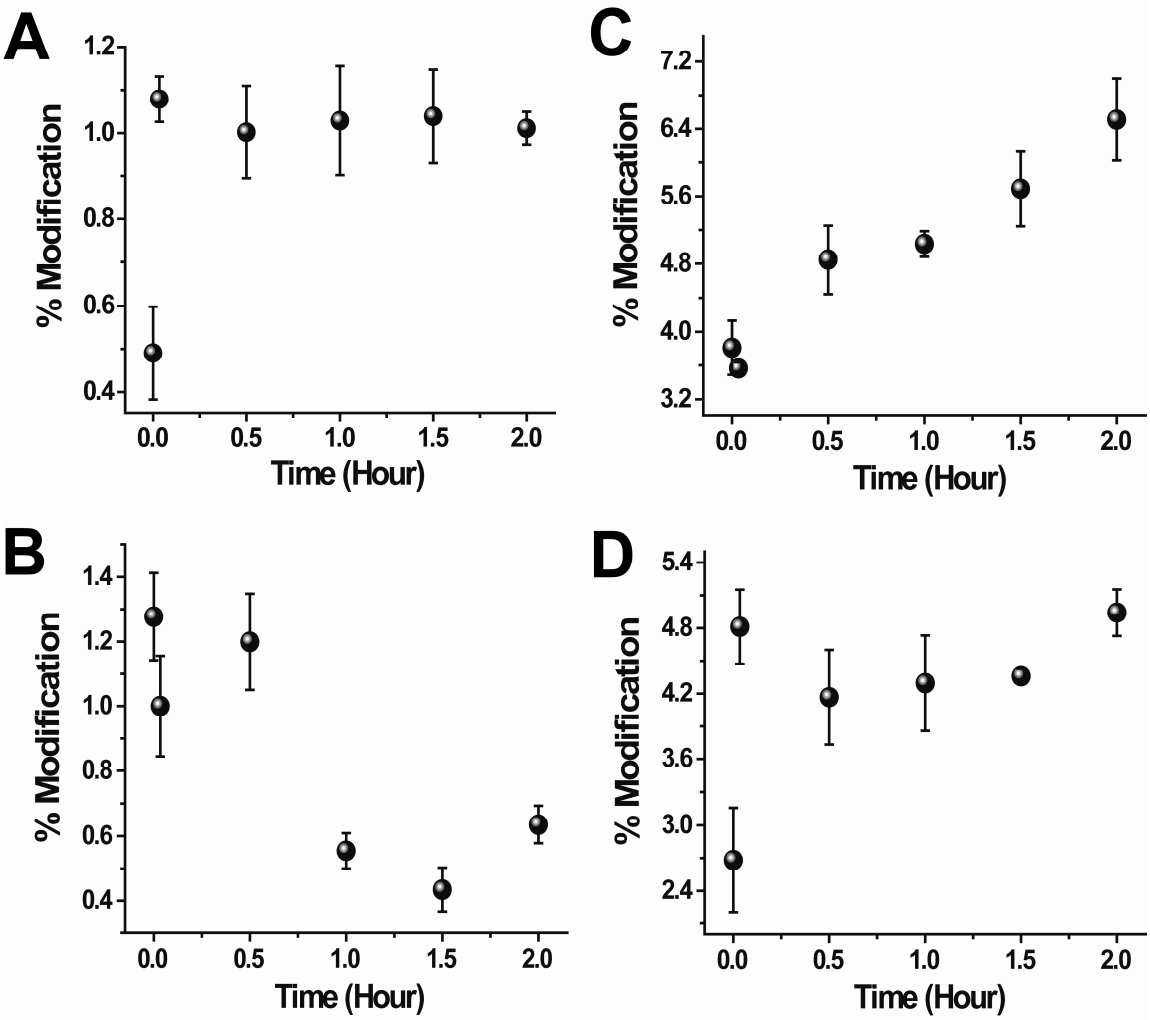


Figure 4.3 Extent of BD modification throughout the course of the dimer formation. (A) Arg3. (B) Arg12. (C) Arg45. (D) Arg97.

A summary of the percentage changes for each of the modified amino acids monitored at t=0 and t=2 hours is shown in Table 4.1.

Table 4.1 Summary of the percentage changes for the modified amino acids from t=0 to t=2 hours after the addition of Cu(II).

Residue	t = 0	t = 2 hours	% Change
NHSA			
N-terminus	84 ± 3	60.7 ± 0.3	-28
Lys6	18.31 ± 0.04	17.9 ± 0.1	-2
Lys19	7.1 ± 0.2	7.2 ± 0.2	0
Lys41/Lys48	13.2 ± 0.7	13.2 ± 0.8	0
Lys58	16.1 ± 0.7	19.0 ± 0.1	+18
Lys75	1.37 ± 0.07	1.62 ± 0.08	+18
Asn83	1.48 ± 0.04	1.46 ± 0.03	0
Lys91	27 ± 1	24.0 ± 0.4	-11
Lys94	3.2 ± 0.3	4.4 ± 0.2	+38
DEPC			
N-terminus	99 ± 3	68.1 ± 0.6	-31
Thr4	86.5 ± 0.7	89.5 ± 0.6	+3
Lys6	8.07 ± 0.05	7.64 ± 0.02	-5
His13	45 ± 1	40.0 ± 0.7	-11
Lys19	13.0 ± 0.9	11.5 ± 0.6	-12
Tyr26	0.74 ± 0.06	0.45 ± 0.06	-39
Ser28	0.24 ± 0.01	0.11 ± 0.01	-54
His31	1.6 ± 0.2	0.8 ± 0.2	-50
Ser33	1.74 ± 0.06	1.19 ± 0.06	-32
Lys41	0.32 ± 0.04	0.34 ± 0.02	0
His51	61 ± 2	66 ± 2	+8
Ser57/Lys58	39 ± 2	47 ± 2	+21
Tyr63	6.3 ± 0.3	3.3 ± 0.2	-48

This table is continued on the next page.

Tyr67	2.2 ± 0.2	1.5 ± 0.1	-32
Lys75	0.31 ± 0.04	0.38 ± 0.02	+23
Ser88	66 ± 2	72 ± 1	+9
Lys94	28 ± 2	36 ± 2	+29
BD			
Arg3	0.5 ± 0.2	1.01 ± 0.04	+100
Arg12	1.3 ± 0.1	0.64 ± 0.06	-51
Arg45	3.8 ± 0.3	6.5 ± 0.5	+71
Arg97	2.7 ± 0.5	4.9 ± 0.2	+81

4.3.4 Determination of Surface Area of β2m Oligomers

ESI-MS experiments were performed to measure the average charge states of β2m oligomers and estimate their surface areas from these measurements using a method recently described by Kaltashov and co-workers [13]. That group recently demonstrated that the average charge state of a protein in its ESI mass spectrum is related to its surface area. Using mass spectral data for six proteins with accurately known surface areas (Table 4.2), we generated a calibration curve by plotting the natural log of the measured average charge state as a function of the natural log of the protein surface area (Figure 4A). The resulting plot has a slope of 0.68 ± 0.02, which compares favorably with the theoretically expected slope of 0.75 and a slope of 0.69 ± 0.02 measured by Kaltashov and co-workers [26]. To obtain the average charge state of each protein, a Gaussian fit of relative abundance of each charge state for the protein in the mass spectrum versus the measured charge states was used (Figure 4B).

Table 4.2 Proteins used to generate the charge-surface area calibration plot.

Protein	PDB ID	Surface Area ($Å^2$)	Average Charge
Ubiquitin	1UBQ	4758	5.5
Cytochrome C	1HRC	6232	7.3
Myoglobin	1WLA	8017	8.2
Ovalbumin	1OVA	15586	14.2
Hemoglobin	1A3N	24548	17.3
Transferrin	1JNF	27395	18.5

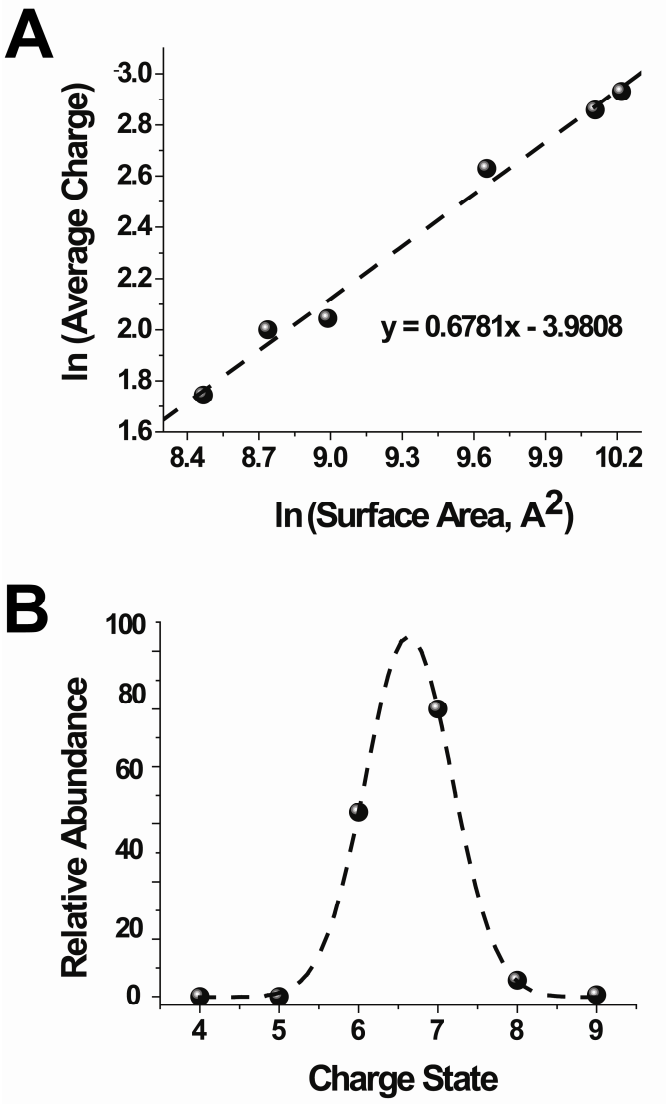


Figure 4.4 (A) Calibration plot obtained from mass spectra of the proteins in Table 4.1, which is used to measure the surface areas of the β2m oligomers. (B) A plot of ion intensity as a function of charge state for the β2m dimer. A Gaussian fit is used to determine the average charge state.

This approach for estimating oligomer surface area was used at different points after adding Cu(II) to initiate the amyloid forming reaction. The average charge states of the monomer and oligomers of β2m ions were determined from ESI mass spectra and

these values were used to estimate their surface areas (Table 4.3). The estimated surface areas of the monomer, dimer, tetramer, and hexamer are 6590 ± 140, 11430 ± 150, 21110 ± 190, and 24960 ± 110 $Å^2$, respectively.

Table 4.3 Measured surface areas of β2m oligomers.

Days of Incubation	Measured Surface Area ($Å^2$)			
	Monomer	Dimer	Tetramer	Hexamer
0	6515	11586	--	--
1	6743	11546	21248	--
3	6655	11435	20843	25050
5	6660	11370	21110	24982
7	6393	11211	21228	24837

4.3.5 Pre-amyloid β2m Dimer

The goal of the current study is to obtain structural information about the β2m dimer that is formed after the protein binds Cu and before it proceeds to form amyloid fibrils. Our group and others have shown that β2m forms dimers, tetramers, and hexamers [1,3]. Because no odd-numbered oligomers are observed, the dimer is likely the basic unit of the tetramer and hexamer. Hence, characterizing the dimeric intermediate will not only help to delineate the early steps of β2m fibrillogenesis but will also aid in determining the structures of the tetramer and hexamer.

Our approach in this work has been to use covalent labeling along with MS detection because of the utility of this method for studying protein-protein interactions in solution. Two mutant β2m oligomer crystal structures are available [10,11], and these

structures provide convenient high resolution information to which our data can be compared. These crystal structures, however, are not of the wild-type protein, and the specific mutations in each case prevent the protein from forming amyloid fibrils. This fact means that our labeling experiments provide more relevant information about the structural changes necessary for wild-type β2m to form amyloid-competent oligomers.

Several factors affect the covalent labeling efficiency of amino acid side chains, but solvent accessibility is the most important [12]. Consequently, covalent labeling is especially effective at identifying residues that mediate protein-protein interactions because the solvent accessibilities of such residues typically undergo noticeable changes. In the case of β2m, complete isolation of the dimer is not possible without some dissociation back to the monomer, so the covalent labeling experiments had to be performed at different times after adding Cu(II) to initiate the amyloid-forming reaction. Our previous work has shown that an appreciable amount of dimer is formed within 30 minutes and no tetramer is formed until after 12 hours. Thus, by monitoring the modification extents of different residues for the first 2 h after the addition of Cu(II), we can obtain some insight into the residues that undergo changes in solvent accessibility over time, suggesting that they are important for mediating dimer formation. As shown in Chapter 3, the β2m residues that react with NHSA, DEPC, and BD are widely distributed along the polypeptide chain and on the surface of the protein. Thus, we cover most regions of the protein. The structural data obtained covers 25 of the protein's 99 residues and ~30% of the surface-exposed residues.

Most of the amino acids probed by the covalent labels undergo a change in reactivity after Cu(II) is added, but we will focus our discussion on those amino acids that undergo changes as the dimer forms in solution. Of the 24 modified residues, only three amino acids, Lys41, Asn83, and Ser88, have no significant change in reactivity after Cu(II) is added and the amyloid-forming reaction has proceeded for 2 h. No change in reactivity suggests that these residues retain a similar chemical environment upon Cu binding and dimer formation. The residues that undergo a change in reactivity upon Cu(II) binding, but change little or none as the dimer forms, are the N-terminus, Arg3, His31, Ser33, Tyr67, Lys91, and Arg97. The changes in reactivity associated with these residues are important because they provide some insight into the structural changes caused by Cu binding that enable the protein to dimerize; however, the focus of this work is the identification of the amino acids that are involved in mediating the interactions between protein units in the dimer. The remaining 14 residues, including Thr4, Lys6, Arg12, His13, Lys19, Tyr26, Ser28, Arg45, His51, Lys58 (and/or Ser57), Tyr63, Lys75, and Lys94, undergo changes in their labeling extent as the dimer's concentration increases in solution. Because solvent accessibility is the key factor controlling the relative reactivity of these residues, the protection from modification of some of these residues in the dimer suggests that these residues are at or near the dimer interface and implies that they may perhaps contribute to the stability of the dimer.

Previous reports showed that the β2m oligomers are native-like in structure [2], so the known structure of the Cu-free monomer is a useful reference point for interpreting our data. In addition, the recent crystal structures of two mutant β2m oligomers, the P32A dimer (PDB ID: 2F8O) [10] and the H13F hexamer (PDB ID: 3CIQ) [11], also provide a

very useful basis for comparison to our data. The symmetry of the H13F hexamer is such that each protein unit interfaces with two other protein units in distinct ways. Thus, together the P32A dimer and the H13F hexamer provide three dimer interfaces to which we can compare our data. The overall fold of β2m changes little in these mutant oligomers, which is consistent with the expected native-like structure of the wild-type oligomers. Upon comparing our data with the mutant oligomers and the wild-type monomer, three important observations are made: (1) dimer formation via D-D strand interactions as seen in the P32A structure and one of the subunit contacts in the H13F structure is improbable; (2) the CFG β-sheet is not part of the dimer interface; and (3) the ABED β-sheet is part of the dimer interface.

Considering the possibility of the D-D β-strand interaction in the dimer, we find that two of the residues expected to be buried by this interaction increase in reactivity as the dimer is formed. The interfaces formed by the interaction of D strands from two monomeric units as seen in the P32A dimer and one of the dimer units (AB chains) in H13F hexamer are shown in Figure 4.5. The two crystal structures show that the interaction surfaces of the anti-parallel D strands span primarily the amino acids Glu50-Lys58. These interfaces predict reduced solvent accessibility of His51 on the D strand and Ser 57 and Lys58 on the DE loop. The calculated SASA of His51 decreases from 93 $Å^2$ in the monomeric protein (PDB ID: 2D4F) to 33 $Å^2$ and 2 $Å^2$ in the P32A dimer and AB dimer in the H13F hexamer, respectively. Similarly, the SASA of Lys58 in the monomer decreases from 214 $Å^2$ to ~167 $Å^2$ in the P32A and 88 $Å^2$ in the AB dimer, respectively. For Ser57, the SASA decreases from 58 $Å^2$ in the monomer to 0 $Å^2$ in the P32A dimer but increases slightly to 65 $Å^2$ in the AB dimer of the H13F hexamer.

Hydrogen-bonding likely occurs between the side chain of His51 and the backbone of Phe56 in the anti-parallel D-D strand interaction of the AB chains of H13F mutant (Figure 4.5B). In addition, in both the P32A and H13F crystal structures, Lys58 can form a salt bridge with Glu50 (Figure 4.5B). These interactions protect residues His51, Ser57, and Lys58 from solvent, and thus, if the dimer interface involves the D strands from two monomers, one would expect these residues to drop in reactivity as the dimer is formed. In our experiments, however, we find that His51, Lys58, and possibly Ser57 increase in reactivity as the dimer's concentration increases, suggesting that these interactions are not present in the wild-type dimer. From studies of the oligomerization properties of a series of β2m mutants in which the four His residues were replaced by different residues, Blaho and Miranker [27] recently speculated that His51 is an important residue at the interface of the tetramer, and therefore not part of the dimer interface. Our results are consistent with this suggestion and provide additional evidence that His51 is not buried upon dimer formation.

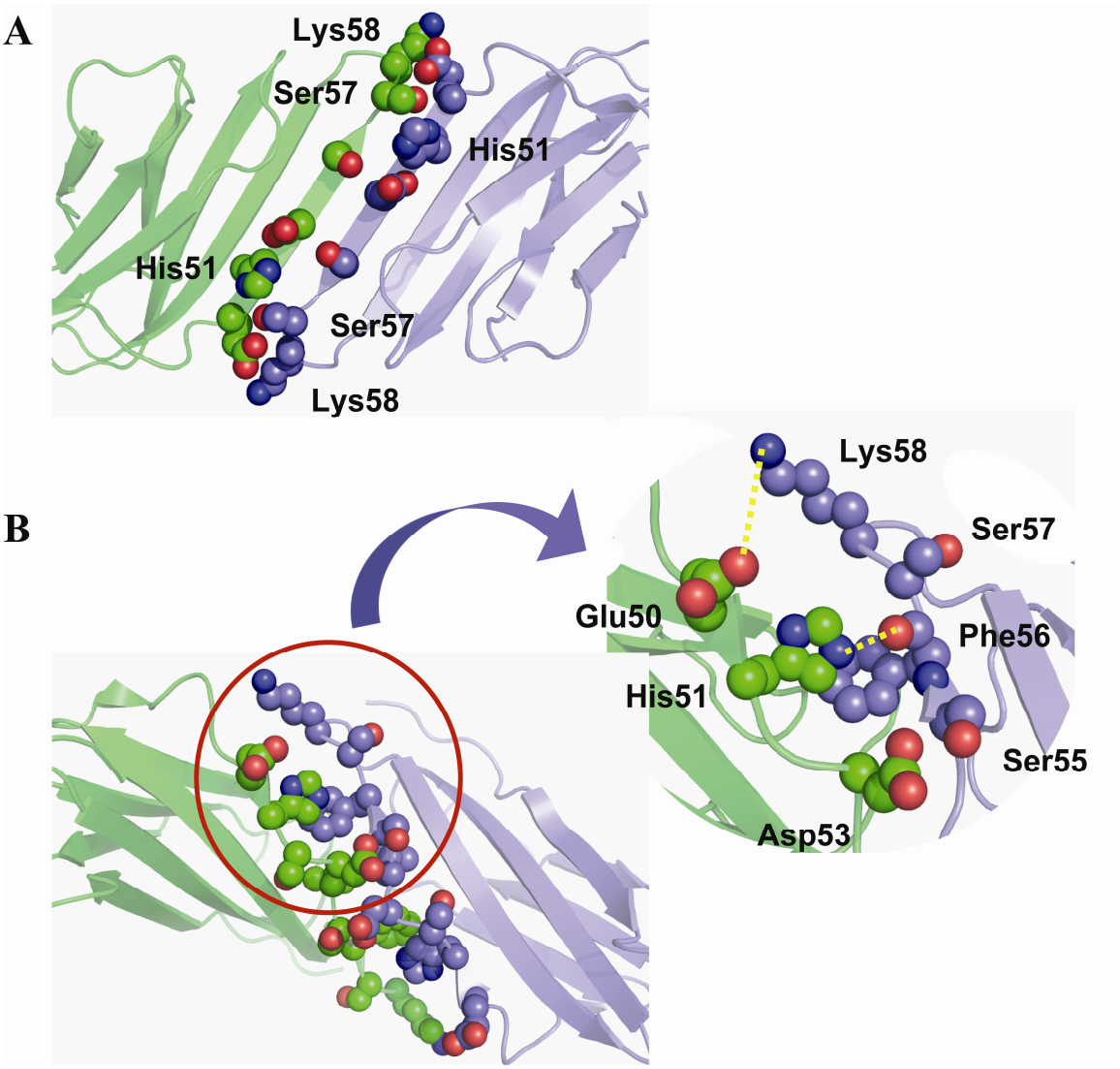


Figure 4.5 Dimers formed by D-D strand interactions. Amino acids Glu50-Lys58 are shown as spheres. (A) D-D strand interface of the crystallographic dimer of P32A. (B) Interaction of adjacent D strands in one of the dimer units (AB chains) in the H13F hexamer. Inter-strand interactions of the side chains of Glu50 and His51 of one monomer with the side chain of Lys58 and the backbone of Phe56, respectively, of another monomer are shown.

The decreased reactivities of other residues provide further evidence against a dimer formed via the D-D strands and give support for a dimer involving the ABED β-sheet. The amino acids that exhibit a decrease in extent of modification upon dimer formation are shown in Figure 4.6A. All of these residues are located on the A, B, and E strands of β2m. Hence, the measured reactivity patterns are consistent with a dimer structure wherein the ABED sheet is part of the dimer interface.

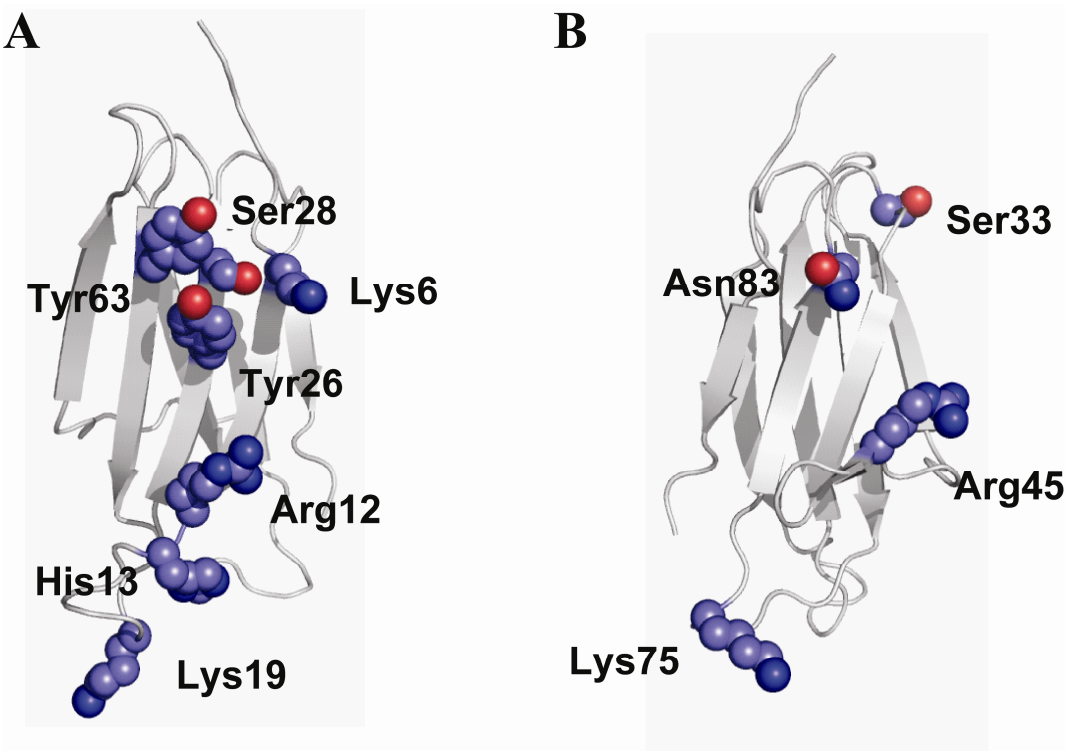


Figure 4.6 The probed amino acids that are located on the (A) ABED and (B) CFG sheets of β2m.

The decreased reactivity of these seven residues are consistent with both an anti-parallel ABED-ABED arrangement as is formed by the B and C chains of the hexameric H13F mutant and a parallel ABED-ABED arrangement (Figure 4.7). Analysis of the BC

dimer from the crystal structure of the hexameric H13F mutant reveals that all residues predicted to occur inside the dimer interface, Lys6 on the A strand, Arg12, His13, and Lys19 on the AB loop, Tyr26 and Ser28 on the B strand, and Tyr63 on the E strand, drop in reactivity after the onset of dimer formation. These data also correlate well with the decrease in calculated SASAs of most of these residues when comparing the monomeric structure to that of the BC dimer of the H13F mutant; Lys6: 199 to 0 $Å^2$; Arg12: 159 to 20 $Å^2$; Lys19: 316 to 90 $Å^2$; Tyr26: 73 to 0 $Å^2$; and Tyr63: 45 to 32 $Å^2$. The SASAs of His13 in the monomer and Phe13 in the BC dimer are 106 $Å^2$ and 0 $Å^2$, respectively. For Ser28, the SASA for monomeric β2m and the BC dimer are both 0 $Å^2$.

Many of these residues form new contacts in the BC chains of the H13F crystal structure that may help stabilize the dimer. For example, Arg12 is very close to Tyr63, and the three carbon chain of Arg's side chain probably forms a hydrophobic interaction with Tyr. Lys19 can form a salt bridge with Asp59 of another monomer in the anti-parallel arrangement (Figure 4.8), which was recently postulated as one of the important interactions caused by Cu binding to monomeric β2m [28]. Indeed, several structural and mutational analyses reveal a crucial role for electrostatic interactions in protein–protein interactions [29,30]. In addition to experimental data, computational studies show that electrostatic interactions can substantially enhance the stability of protein-protein complexes [31,32]. Lys19 is also in close proximity to Glu16, indicating that Lys19 might form a complex salt bridge with Glu16 of the same protein unit and Asp59 of another protein unit. Such networks involving two or more ion pairs are common at protein-protein interfaces and contribute more to protein stability than isolated ion pairs, which predominate in intramolecular interactions [33-35].

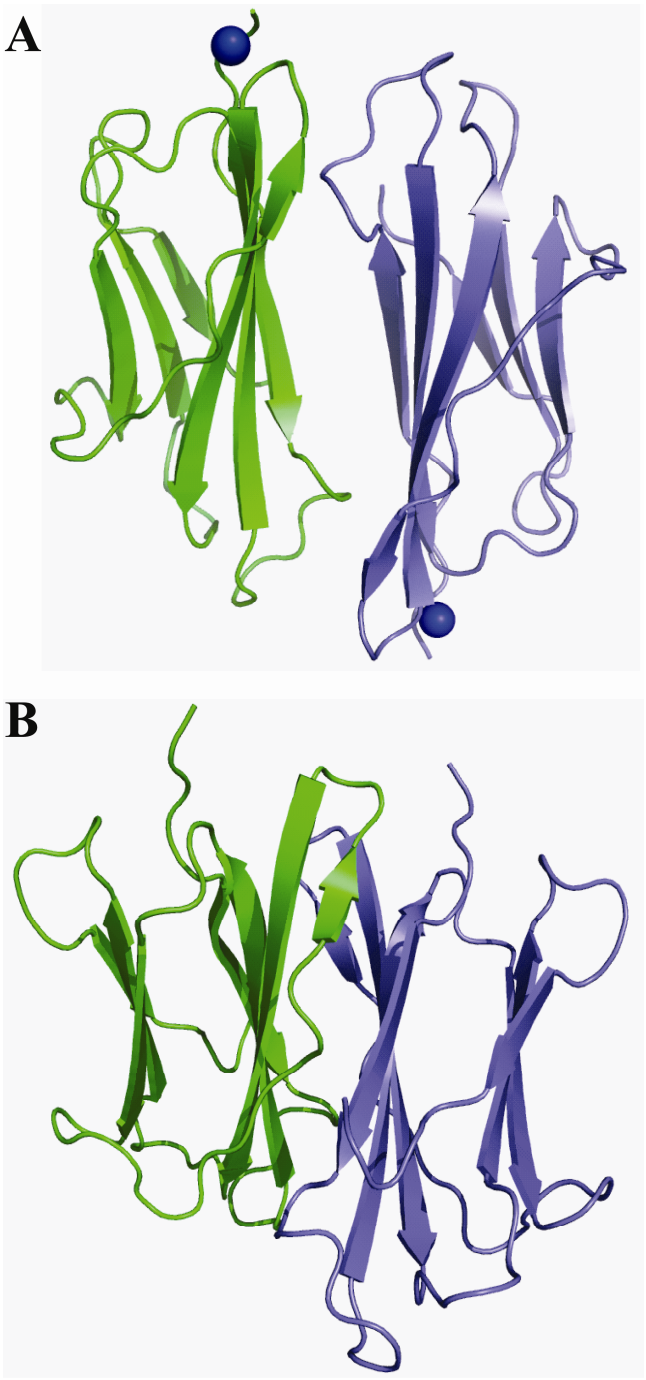


Figure 4.7 Two possible dimers formed via ABED-ABED inter-sheet interactions. (A) Anti-parallel ABED-ABED arrangement as seen in the dimer formed by the B and C chains of the H13F mutant (PDB ID: 3CIQ). (B) Parallel ABED-ABED arrangement as calculated by the docking of two energy-minimized Cu-free β2m monomers obtained from the MHC complex (PDB ID: 1DUZ).

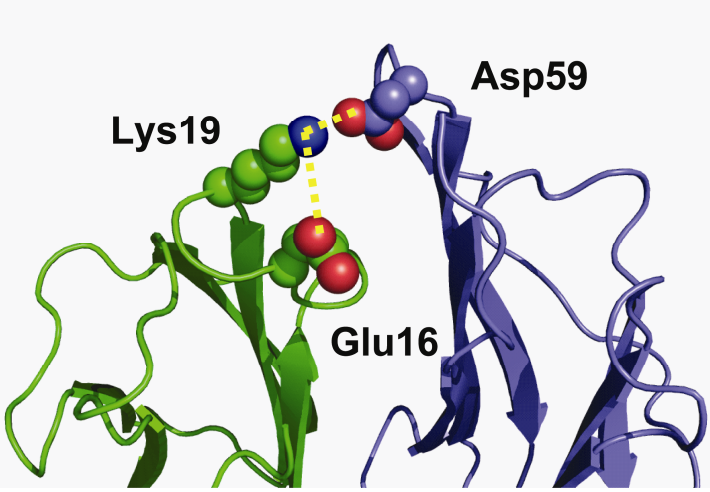


Figure 4.8 Interface between the chains B and C in an anti-parallel arrangement in the H13F hexamer (PDB ID: 3CIQ), showing the inter-subunit salt bridges involving Asp59 and Lys19 and the complex salt bridge involving Glu16.

Detailed structural information is not available for a dimer with a parallel ABED-ABED interface, but we have used protein docking to test the viability of this potential dimer interface. Docking calculations involving two β2m monomers reveal that the dimer with the highest docking score has a parallel ABED-ABED interface with a crossing angle of ~35^{o} (Figure 4.7B). Analysis of this docked dimer structure indicates that the residues that experimentally decrease in reactivity are also part of the dimer interface, with the exception of Lys19. The calculated SASAs for the residues in the docked dimer structure when compared to the monomeric crystal structure are: Lys6: 199 to 19 Å^2; Arg12: 159 to 0 Å^2; His13: 106 to 78 32 Å^2; Lys19: 316 to 194 Å^2; Tyr26: 73 to 0 Å^2; Ser 28: 0 to 0 Å^2; and Tyr63: 45 to 0 Å^2. The decrease in SASA of Lys19 in the parallel dimer structure may be a consequence of forming a salt bridge with Glu16 of the same monomeric unit. Even though the drop in SASAs of most of these residues is consistent with the experimentally determined reactivities, close inspection of this docked dimer

structure does not reveal any favorable salt bridge interactions between the two monomeric units. Indeed, calculations of the electrostatic component of the binding free energy indicate that the parallel arrangement has a slightly unfavorable electrostatic component. Because our previous work demonstrated the importance of salt bridges in dimer formation [30], we conclude that this parallel ABED-ABED interface is less likely than the anti-parallel one.

The anti-parallel ABED-ABED arrangement is also consistent with the measured surface area of the β2m dimer. The measured surface area of the monomer (6590 ± 140 $Å^2$) compares well with the surface area (6500 $Å^2$) obtained from the monomeric crystal structure of β2m. The surface areas of the crystallographic dimer of the P32A dimer, the AB dimer of the H13F mutant, the BC dimer of the H13F mutant, and the parallel ABED-ABED arrangement from the docking calculations are 11988 $Å^2$, 12093 $Å^2$, 11524 $Å^2$, and 11095 $Å^2$, respectively. The measured surface area for the dimer (11430 ± 150 $Å^2$) compares most favorably with the surface area of the anti-parallel ABED-ABED arrangement as seen in the BC chains of the H13F mutant.

There are two other possible dimer forms with an ABED β-sheet interaction: (1) a parallel 3-stranded (CFG) and 4-stranded (ABED) interaction, and (2) an anti-parallel CFG-ABED interaction. Only four residues in the CFG sheet are probed by the covalent labels: His31, Arg45, Lys75, and Asn83 (Figure 4.7B). If the inter-sheet interaction involves the CFG sheet together with an ABED sheet, these four residues are expected to decrease in reactivity as the dimer forms. Our data clearly shows, though, that the reactivity of Arg45 on the CD loop and Lys75 on EF loop increase upon dimer formation indicating that these residues are more solvent accessible. There is no significant change

in the reactivity of Asn83 on the F strand, suggesting that this residue is not part of the dimer interface. His31 on the BC loop had an initial drop in reactivity after the addition of Cu(II), but its reactivity remains constant during dimer formation. Hence, the reactivity pattern of these residues suggests that a dimer formed by the interaction of a CFG from one protein unit with an ABED sheet of another protein unit is unlikely.

Overall, our data support a β2m dimer interface that involves an antiparallel ABED-ABED interaction. As described above, the seven residues that are found at this interface all decrease in reactivity upon dimer formation. There are, however, seven other residues, Thr4, Arg45, His51, Ser57, Lys58, Lys75, and Lys94, that increase in reactivity as the dimer's concentration increases in solution. Are the reactivities of these residues consistent with the proposed dimer structure? The increased reactivity of five of these residues can be rationalized by comparing the monomeric structure and the crystal structure of the BC chains of the H13F hexamer. The increased reactivity of Arg45 might be explained by the breaking of a salt bridge interaction with Asp38 that is probably present in the monomer but not in the BC dimer of the H13F mutant. Arg45's SASA also increases slightly from 109 to 118 $Å^2$. Similarly, Lys94 probably forms a salt bridge with Glu77 in the monomer, but these two residues are much further apart in the BC dimer of the H13F mutant. The increased reactivities of His51, Lys58 and Lys75 correlate reasonably well with the increase in SASAs of these residues in the BC chains of the H13F mutant relative to the monomer: His51: 93 to 113 $Å^2$; Lys58: 214 to 327 $Å^2$; and Lys75: 245 to 302 $Å^2$.

4.4 Conclusions

In conclusion, our covalent labeling data indicate that the β2m dimer, which is formed prior to the β2m amyloid fibrils, has an interface that involves the antiparallel arrangement of ABED sheets from two monomers. Our confidence in this assignment comes from labeling data that covers almost one-half of the amino acids that are at the interface of the dimer and from comparison to the crystal structure of the hexameric H13F mutant, which appears to include many of the interactions present in the wild-type dimer. Our data clearly indicate that the dimer interface present in the P32A mutant of β2m is not the interface present in the wild-type dimer. Because β2m also forms tetramers and hexamers and not odd-ordered oligomers, the knowledge gained in this study about the dimer interface will serve as important foundational information for characterizing these larger oligomers. Information about these oligomeric intermediates should be useful for the development of therapeutics against DRA. More broadly speaking, covalent labeling along with MS detection appears to be a useful method for probing protein–protein interactions in situations where more than one protein species is present in solution. This ability could make this method useful for other amyloid forming systems in which discrete oligomers precede fibril formation. For some amyloid forming proteins, these small prefibrillar oligomers are thought to be responsible for cellular toxicity rather than the amyloid fibrils themselves [8,9], indicating the importance of studying their structures.

4.5 References

1. Calabrese, M.F., and Miranker, A.D. (2007) Formation of a stable oligomer of β-2 microglobulin requires only a transient encounter with Cu(II). *J. Mol. Biol. 367*, 1-7.

2. Eakin, C.M., Attenello, F.J., Morgan, C.J., and Miranker, A.D. (2004) Oligomeric assembly of native-like precursors precedes amyloid formation by β-2 microglobulin. *Biochemistry 43*, 7808-7815.

3. Antwi, K., Mahar, M., Tyson, J.F., and Vachet, R.W. (2007) Cu(II) organizes β-2-microglobulin oligomers but is released before amyloid formation. *Protein Sci. 17*, 748-759.

4. Walsh, D.M., Lomakin, A., Benedek, G.B., Condron, M.M., and Teplow, D.B. (1997) Amyloid beta-protein fibrillogenesis – detection of a protofibrillar intermediate. *J. Biol. Chem. 272*, 22364-22372.

5. Friedhoff, P., von Bergen, M., Mandelkow, E.M., Davies, P., and Mandelkow, E. (1998) A nucleated assembly mechanism of Alzheimer paired helical filaments. *Proc. Natl. Acad. Sci. USA 95*, 15712-15717.

6. Ferrao-Gonzales, A.D., Robbs, B.K., Moreau, V.H., Ferreira, A., Juliano, L., Valente, A.P., Almeida, F.C.L., Silva, J.L., and Foguel, D. (2005) Controlling beta-amyloid oligomerization by the use of naphthalene sulfonates – trapping low molecular weight oligomeric species. *J. Biol. Chem. 280*, 34747-34754.

7. Smith, A.M., Jahn, T.R., Ashcroft, A.E., and Radford, S.E. (2006) Direct observation of oligomeric species formed in the early stages of amyloid fibril formation using electrospray ionization mass spectrometry. *J. Mol. Biol. 364*, 9-19.

8. Bucciantini, M., Giannoni, E., Fabrizio, C., Baroni, F., Formigli, L., Zurdo, J., Taddei, N., Ramponi, G., Dobson, C.M., and Stefani, M. (2002) Inherent toxicity of aggregates implies a common mechanism for protein misfolding diseases. *Nature 416*, 507-511.

9. Mendes Sousa, M., Cardoso, I., Fernandes, R., Guimaraes, A., and Saraiva, M.J. (2001) Deposition of transthyretin in early stages of familial amyloidotic polyneuropathy. Evidence for toxicity of nonfibrillar aggregates. *Am. J. Pathol. 159*, 1993-2000.

10. Eakin, C. M., Berman, A. J., and Miranker, A. D. (2006) A native to amyloidogenic transition regulated by a backbone trigger. *Nature Struct. Mol. Biol. 13*, 202-208.

11. Calabrese, M. F., Eakin, C. M., Wang, J. M., and Miranker, A. D. (2008) A regulatable switch mediates self-association in an immunoglobulin fold. *Nature Struct. Mol. Biol. 15*, 965-971.

12. Mendoza, V. L., and Vachet, R. W. (2009) Probing protein structure by amino acid-specific covalent labeling and mass spectrometry. *Mass Spec. Rev. 28*, 785-815.

13. Kaltashov, I. A., and Mohimen, A. (2005) Estimates of protein surface areas in solution by electrospray ionization mass spectrometry. *Anal. Chem. 77*, 5370-5379.

14. Kihara, M., Chatani, E., Iwata, K., Yamamoto, K., Matsuura, T., Nakagawa, A., Naiki, H., and Goto, Y. (2006) Conformation of amyloid fibrils of β2-microglobulin probed by tryptophan mutagenesis. *J. Biol. Chem. 281*, 31061-31069.

15. Fraczkiewicz, R., and Braun, W. (1998) Exact and Efficient analytical calculation of the accessible surface areas and their gradients for macromolecules. *J. Comp. Chem. 19*, 319-333.

16. Khan, A. R., Baker, B. M., Ghosh, P., Biddison, W. E., and Wiley, D. C. (2000) The structure and stability of an HLA-A*0201/octameric tax peptide complex with an empty conserved peptide-N-terminal binding site. *J. Immunol. 164*, 6398-6405.

17. Verdone, G., Corazza, A., Viglino, P., Pettirossi, F., Giorgetti, S., Mangione, P., Andreola, A., Stoppini, M., Bellotti, V., and Esposito, G. (2002) The solution structure of human β2-microglobulin reveals the prodromes of its amyloid transition. *Protein Sci. 11*, 487-499.

18. Mahoney, M. Jorgensen, A. (2000) A five-site model for liquid water and the reproduction of the density anomaly by rigid, nonpolarizable potential functions. *J. Chem. Phys. 112*, 8911-8922.

19. Cornell, W., Cieplak, P., Bayly, C., Gould, I., Merz, K., Ferguson, D., Spellmeyer, D., Fox, T., Caldwell, J., and Kollma, P. (1995) A second generation force field for the simulation of proteins, nucleic acids, and organic molecules. *J. Am. Chem. Soc. 117*, 5179-5197.

20. Sorin, E. J. and Pande, V. S. (2005) Exploring the helix-coil transition via all-atom equilibrium ensemble simulations. *Biophys. J. 88*, 2472-2493.

21. van der Spoel, D., Lindahl, E., Hess, B., Groenhof, G., Mark, A., and Berendsen, H. (2005) Gromacs: Fast, flexible, and free. *J. Comp. Chem. 26*, 1701-1719.

22. Lindahl, E., Hess, B., and van der Spoel, D. (2001) Gromacs 3.0: A package for molecular simulation and trajectory analysis. *J. Mol. Mod. 7*, 306-317.

23. Berendsen, H., van der Spoel, D., and van Drunen, R. (1995) Gromacs: A message-passing parallel molecular dynamics implementation. *Comp. Phys. Comm. 91*, 43-56.

24. Chen, T., Li, L., Weng, Z. (2002) Docking unbound proteins using shape complementary, desolvation, and electrostatics. *Proteins 47*, 281-294.

25. Chen, T. and Weng, Z. (2003) ZDOCK: an initial-stage protein-docking algorithm. *Proteins 52*, 80-87.

26. Holst, M. J., Baker, N. A., and Wang, F. (2000) Adaptive multilevel finite element solution of the Poisson-Boltzmann equation. I. Algorithms and examples. *J. Comput. Chem. 21*, 1319-1242.

27. Blaho, D. V., and Miranker, A. D. (2009) Delineating the conformational elements responsible for Cu2+-induced oligomerization of beta-2 microglobulin. *Biochemistry 48*, 6610-6617.

28. Srikanth, R., Mendoza, V. L., Bridgewater, J. D., Zhang, G., and Vachet, R. W. (2009) Copper binding to β-2-microglobulin and its pre-amyloid oligomers. *Biochemistry 48*, 9871-9881.

29. Sheinerman, F. B., Norel, R., and Honig, B. (2000) Electrostatic aspects of protein-protein interactions. *Curr. Opin. Struct. Biol. 10*,153–159.

30. Keskin, O., Gursov, A., Ma, B., and Nussinov, R. (2008) Principles of protein-protein interactions: What are the preferred ways for proteins to interact? *Chem. Rev. 108*, 1225-1244.

31. Kundrotas, P. J., and Alexov, E. (2006) Electrostatic Properties of Protein-Protein Complexes. *Biophys. J. 91*, 1724–1736.

32. Dong, F., and Zhou, H.-X. (2006) Electrostatic Contribution to the Binding Stability of Protein–Protein Complexes. *Proteins 65*, 87–102.

33. Musafia, B., Buchner, V., and Arad, D. (1995) Complex salt bridges in proteins: Statistical analysis of structure and function. *J Mol. Biol. 254*, 761-770.

34. Risal, D., Gourinath, S., Himmel, D. M., Szent-Gyorgyi, A. G., and Cohen, C. (2004) Myosin subfragment 1 structures reveal a partially bound nucleotide and a complex salt bridge that helps couple nucleotide and actin binding. *Proc. Natl. Acad. Sci. USA 101*, 8930-8935.

35. Xu, G. Z., Liu, R. T., Zak, O., Aisen, P., and Chance, M. R. (2005). Structural allostery and binding of the transferrin·receptor complex. *Mol. Cell. Proteomics 4*, 1959-1967.

CHAPTER 5

STRUCTURE OF THE PRE-AMYLOID TETRAMER OF β-2-MICROGLOBULIN FROM COVALENT LABELING AND MASS SPECTROMETRY

This chapter is part of a paper in preparation: Mendoza, V.L., Baron-Rodriguez, M.A., Blanco, C., and Vachet, R.W. (2010) Structure of the pre-amyloid tetramer of β-2-microglobulin from covalent labeling and mass spectrometry.

5.1 Introduction

The formation of discrete oligomers as prefibrillar intermediates has been observed for several proteins including β2m [1-5]. The fact that amyloidogenic proteins form oligomeric intermediates prior to fibril formation may be important because the oligomers themselves might be therapeutic targets for preventing amyloid fibril formation Indeed, recent studies of other amyloid systems suggest that prefibrillar intermediates might be responsible for cellular toxicity rather than the amyloid fibrils [6,7].

Our previous work has shown that Cu(II)-induced β2m amyloid formation is preceded by the formation of discrete, oligomeric intermediates, including dimers, tetramers, and hexamers [4]. In our proposed model, monomeric β2m binds Cu(II) and is destabilized. Cu(II)-bound β2m forms a dimer which subsequently forms tetramer I. Tetramer I undergoes a conformational change, releases Cu(II), and undergoes rearrangement to form tetramer II. Tetramer II associates with a dimer that presumably releases its bound Cu(II) and forms the hexamer. Several studies have attempted to characterize these pre-amyloid oligomers, but this task is very challenging because the oligomers are present as a mixture of species and, as intermediates, are only transiently

populated. One successful approach to obtain structural information about the oligomers has been to create β2m mutants that are stable enough to crystallize as oligomers [8,9]. In this way, it has been found that a P32A mutant of β2m forms a dimer in the absence of Cu(II) and an H13F mutant forms a hexamer in the presence of Cu(II). These crystal structures provide high resolution atomic-level information about possible β2m oligomer structures, but they do not provide a complete picture about oligomer structure because these mutants do not ultimately form amyloid fibrils.

As a complement to these crystallographic studies, covalent labeling with mass spectrometric method is used to study the oligomeric intermediates. We recently used this method to gain insight into the structure of the dimer formed by wild-type β2m [10] and found that the dimer interface is formed by the anti-parallel stacking of ABED β-sheets from two β2m monomers. In this chapter, we applied this covalent labeling/MS approach to characterize the tetrameric form of β2m that follows formation of the dimer. Using the covalent labeling data along with molecular dynamics calculations, we are able to then build a model of the tetramer that indicates how the protein can continue to form higher order oligomers.

5.2 Experimental Procedure

5.2.1 Materials

The same materials were used as described in Section 3.2.1.

5.2.2 Formation of β2m Oligomers

The formation of β2m oligomers was done in a manner identical to that described in Section 4.2.2.

5.2.3 Covalent Labeling Reactions

The covalent labeling reactions were done in a manner identical to that described in Sections 3.2.3-3.2.5.

5.2.4 Proteolytic Digestion

The proteolytic digestions were performed in a manner identical to that described in Section 3.2.6.

5.2.5 Instrumentation

The amount of modification was determined by removing an aliquot of the purified β2m and analyzing the samples using a Bruker Amazon quadrupole ion trap mass spectrometer (Billerica, MA) equipped with an ESI source. The ESI source was operated at a spray voltage of 4.5 kV, and the capillary temperature was set at 220°C. The voltages for the transfer optics between the ESI source and the ion trap were optimized for maximum signal.

The proteolytic fragments were separated by an Agilent HP1100 (Wilmington, DE) HPLC system with a C18 column (15 cm x 2.1 mm, 5 μm particle size, Supelco, St. Louis, MO) for on-line analysis by MS and MS/MS. The fragments of NHSA- and BD-modified β2m were eluted using a linear gradient of methanol that increased from 5 to

70% over 20 min and 70 to 100% over the final 2 min at a flow rate of 0.250 mL/min. For the fragments of DEPC-modified β2m, a linear gradient of methanol that increased from 5 to 70% over 30 min and 70 to 100% over the final 3 min was used. For both gradient conditions, water comprised the balance of the solvent, and a total of 0.1% acetic acid was present. The LC effluent was fed into the mass spectrometer with similar ESI source conditions as described above. Tandem mass spectra were acquired using collision induced dissociation (CID) with isolation widths of 1.0 Da and excitation voltages between 0.6 and 1.0 V.

5.2.6 Amino Acid Modification Percentage

The amino acid modification percentages were calculated in a manner identical to that described in Section 3.2.8.

5.2.7 Molecular Dynamics Simulations

All molecular dynamics simulations were initiated using a dimer, the B and C chains, from the H13F mutant hexamer crystal structure (PDB ID: 3CIQ) [9], which contains 98 amino acid residues. Met0 was eliminated from the original set of coordinates. Two dimers were added together before the simulations began in order to produce a tetramer structure that was consistent with the covalent labeling data. This tetramer structure was then energy-minimized, using an explicit model of the solvent, with the AMBER94 force field [11,12] implemented in GROMACS [13-15]. During the energy minimization hydrogen atoms were explicitly added to the tetramer model and His residues were considered to be neutral, as shown in nuclear magnetic resonance (NMR)

studies by Esposito and co-workers [16]. Water molecules were simulated using a TIP3P model [17] and molecules were added until a cubic simulation cell of side length 9.6 nm was achieved.

5.3 Results and Discussion

Structural studies of the oligomeric intermediates that precede β2m amyloid formation are challenging because these oligomers are transient, present as a mixture, and can dissociate back into the monomer during isolation. We reported previously that β2m oligomers are formed via the discrete addition of dimer units [4]. The dimer first appears within 1 hour after adding Cu(II), a first form of the tetramer appears within 12-24 hours, a second Cu(II)-free form of the tetramer appears after approximately 3 days, and the hexamer appears soon after the Cu(II)-free tetramer. Since complete isolation of the tetramer is not possible without some dissociation back to the dimer and monomer, covalent labeling experiments were performed at different times before (t = 0 min) and after (t = 2 min, 2 hrs, 0.5 day, 1 day, 1.5 days, 2 days, and 2.5 days) adding Cu(II) to initiate the amyloid fibril-forming reaction. At time points between 0.5 and 2.5 days, the mixture contains monomer, dimer, and the first form of the tetramer. The percentage of tetramer at time points after the addition of Cu(II) increases from 0 to 15% based on ESI-MS measurements of the desalted sample under amyloid-forming conditions, as described previously [4]. Also, the percentage of dimer increases from 0 to 28% from 0 to 2.5 days. In recent covalent labeling studies of the β2m dimer, we identified 20 amino acids that undergo changes in labeling reactivity, relative to the monomer, after Cu(II) binding and dimer formation (Figure 5.1) [10].

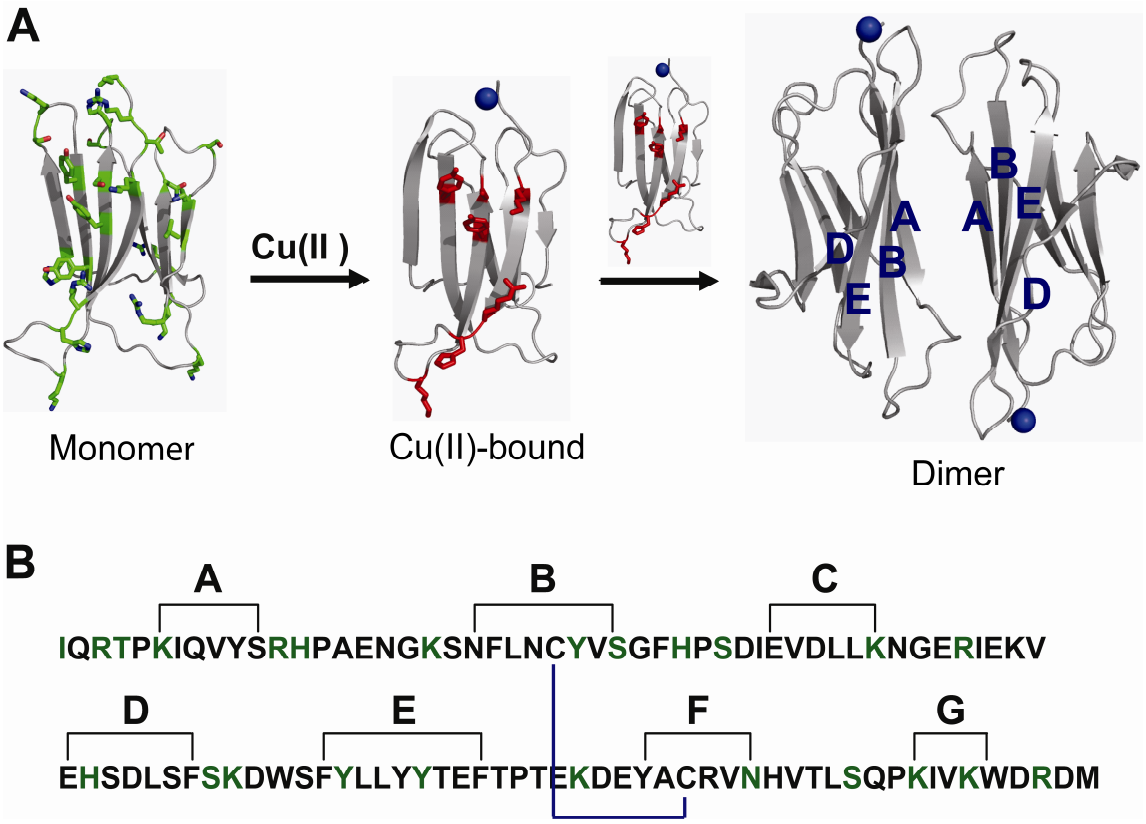


Figure 5.1 (A) Ribbon representation of monomeric β2m (PDB ID: 2D4F), Cu(II)-bound β2m, and the formation of the dimer by stacking of two antiparallel ABED sheets. Amino acids modified by the covalent labels are shown as green sticks. The amino acids found in the dimer interface (ABED β-sheet) are shown as red sticks. (B) Amino acid sequence of β2m showing strand nomenclature [18]. Black lines show amino acids on each β strand. The internal disulfide bond is shown in blue. The amino acids probed by the covalent labels are shown in green.

In the current study, we sought to identify residues that only undergo changes in labeling reactivity as the tetramer is formed. To do this the covalent labeling data that is obtained when the tetramer is present (0.5-2.5 days) is compared to the data acquired when only dimer and monomer are present in solution (0.5-2 hours). Any residues that undergo new changes in reactivity after 0.5 days are then residues that undergo changes in solvent accessibility, and thus are likely important for the formation of the tetramer.

5.3.1 Covalent Labeling with NHSA

Amino groups such as the ε-NH_2 of lysine residues and the N-terminal α-NH_2 can react with NHSA, but information from such surface mapping experiments is reliable only if the structural integrity of a protein is preserved during the reaction. Previously, we have reported that reaction of β2m with a 3-fold molar excess of NHSA for 1 min at 37 ^{o}C does not induce any protein structural changes [10]. The reaction time was kept short (1 min) to avoid tetramer dissociation.

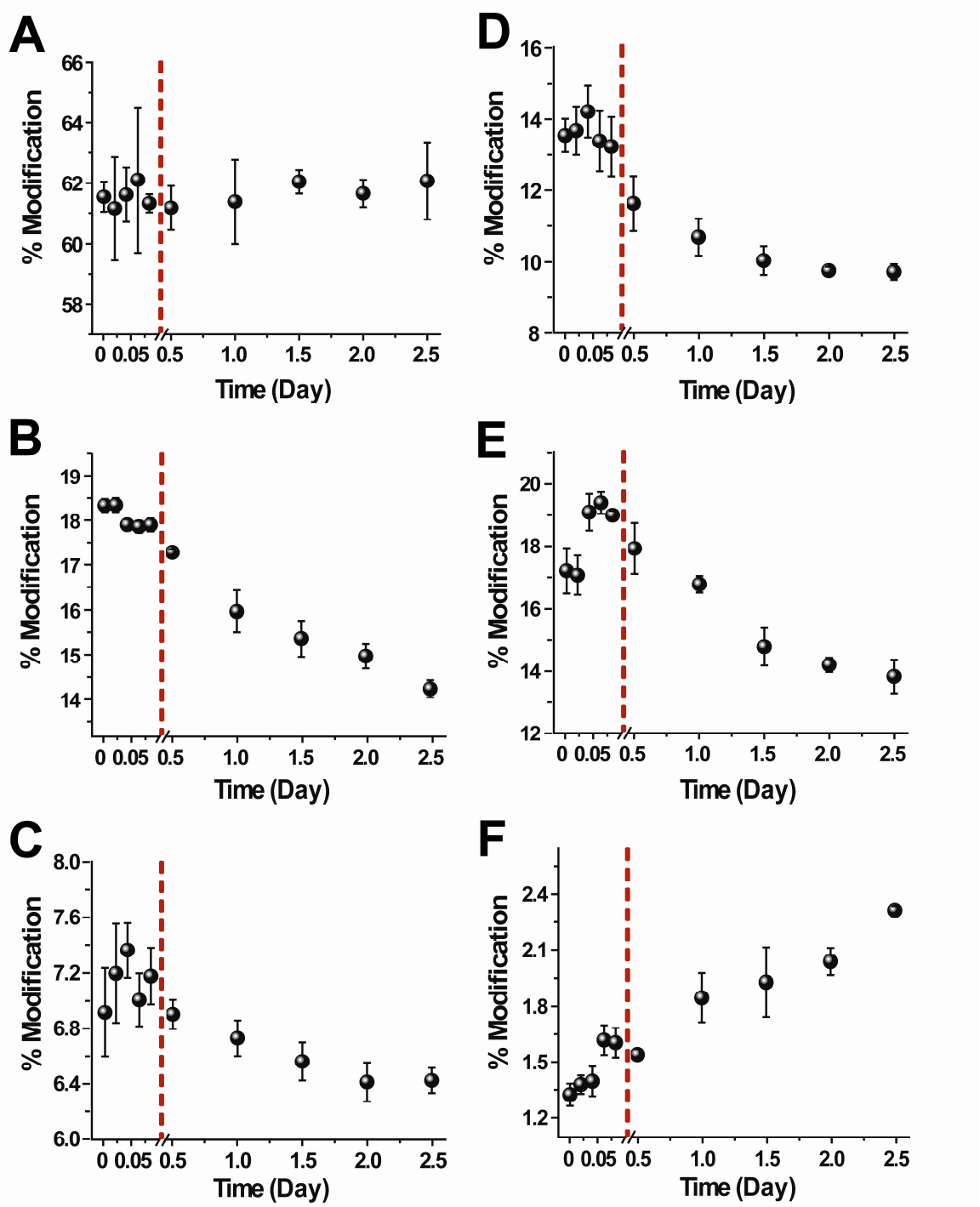
Proteolytic digestion and LC-MS analyses are necessary to identify the amino acids that change in reactivity as the tetramer concentration increases. Under solution conditions in which β2m forms amyloid fibrils, LC-MS/MS analyses show that the N-terminus, almost all of the lysines (Lys6, Lys19, Lys41, Lys58, Lys75, Lys91, and Lys94), and Asn83 are labeled to different degrees. The unmodified and modified fragments containing these amino acids are detectable at all time points; however, the level of modification for some residues changes over time (*vide infra*). In our previous study, definitive MS/MS data could not be obtained for the fragment Lys41-Leu54 [10]. Here, the better sensitivity of the quadrupole ion trap used in these experiments allowed identification of Lys41 as the modification site.

The reactions with NHSA at different time points after formation of the tetramer reveal that the modification extents of some residues change as the concentration of the tetramer increases in solution (Figure 5.2). The percentage modification for the peptide fragments containing the indicated residues were determined from the LC-MS intensities of the modified and unmodified fragments, as described in the Materials and Methods section. For each time point, the modification reaction was repeated three times. Because

ion intensity ratios of the unmodified and modified peptide fragments are used to determine the extent of modification, small changes in modification levels can be accurately and precisely determined, as we demonstrated in our previous work [10].

Of all the residues acetylated by NHSA, only the N-terminus (Figure 5.2A) does not change in the extent of modification from 0.5 to 2.5 days after addition of Cu(II). The levels of modification of Lys75 (Figure 5.2F) and Asn83 (Figure 5.2G) increase from 0.5 days after Cu(II) is added to 2.5 days after the metal is added. In contrast, the reactivity of Lys6, Lys19, Lys41, Lys58, Lys91, and Lys94 clearly decrease as the tetramer's concentration increases in solution. The extent of modification of Lys6 decreases slightly within 2 hours after adding Cu(II), and the reactivity of this residue continues to decrease up to 2.5 days. In our previous covalent labeling studies on the dimer, we concluded that Lys6 was part of the dimer interface, and so its continual decrease in reactivity is consistent with the dimer's concentration continuing to increase in solution over time.

The rest of the lysine residues, Lys19, Lys41, Lys58, Lys91, and Lys94, behave differently. The reactivity of these five lysine side chains also decrease as the tetramer concentration increases, but their reactivities either do not change or increase prior to tetramer formation. For example, Lys94 increases in reactivity as the dimer's concentration increases in solution, but as soon as the tetramer begins to significantly populate the solution, Lys94 decreases in reactivity. On the other hand, Lys41 undergoes no change in reactivity as the dimer is formed, whereas it decreases in reactivity as the tetramer's concentration increases in solution.



This figure is continued on the next page.

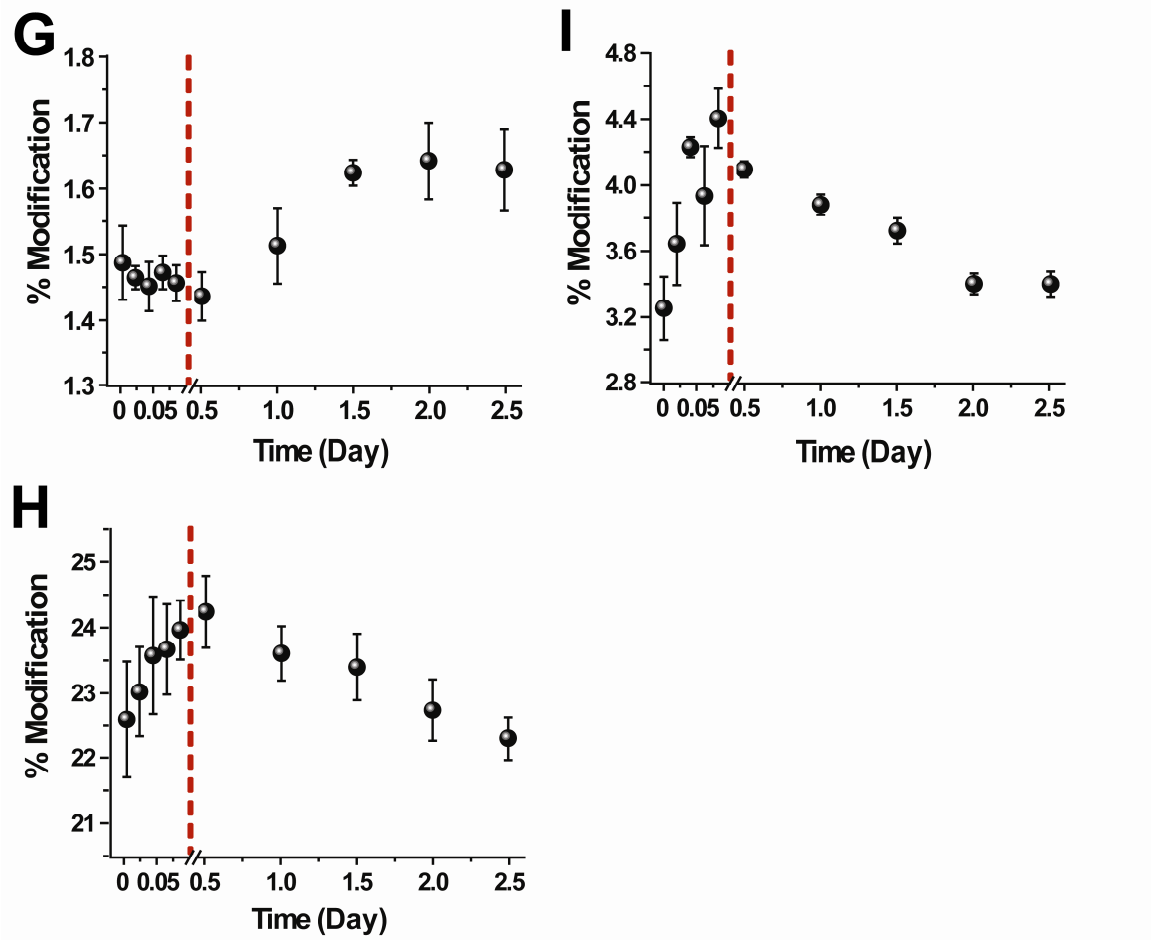


Figure 5.2 Extent of NHSA modification throughout the course of tetramer formation. (A) N-terminus. (B) Lys6. (C) Lys19. (D) Lys41. (E) Lys58. (F) Lys75. (G) Ans83. (H) Lys91. (I) Lys94. The changes in modification during the dimer formation are shown (points before red line) as reference points.

In summary, our results indicate that the reactivity of NHSA with the N-terminus does not change as the tetramer is formed in solution. In contrast, the reactivities of Lys75 and Asn83 increase as the tetramer is formed, whereas the reactivities of Lys6, Lys19, Lys41, Lys58, Lys91, and Lys94 decrease as the tetramer's concentration increases in solution. Importantly, the modification extent for Lys6 decreases within the first 2 h after adding Cu(II), which suggests that its continuing drop in modification

extent is due to the formation of more dimer in solution rather than the formation of tetramer. Tetramer formation is then likely responsible for the decreased reactivity of the remaining lysine residues (i.e. Lys19, Lys41, Lys58, Lys91, and Lys94).

5.3.2 Covalent Labeling with DEPC

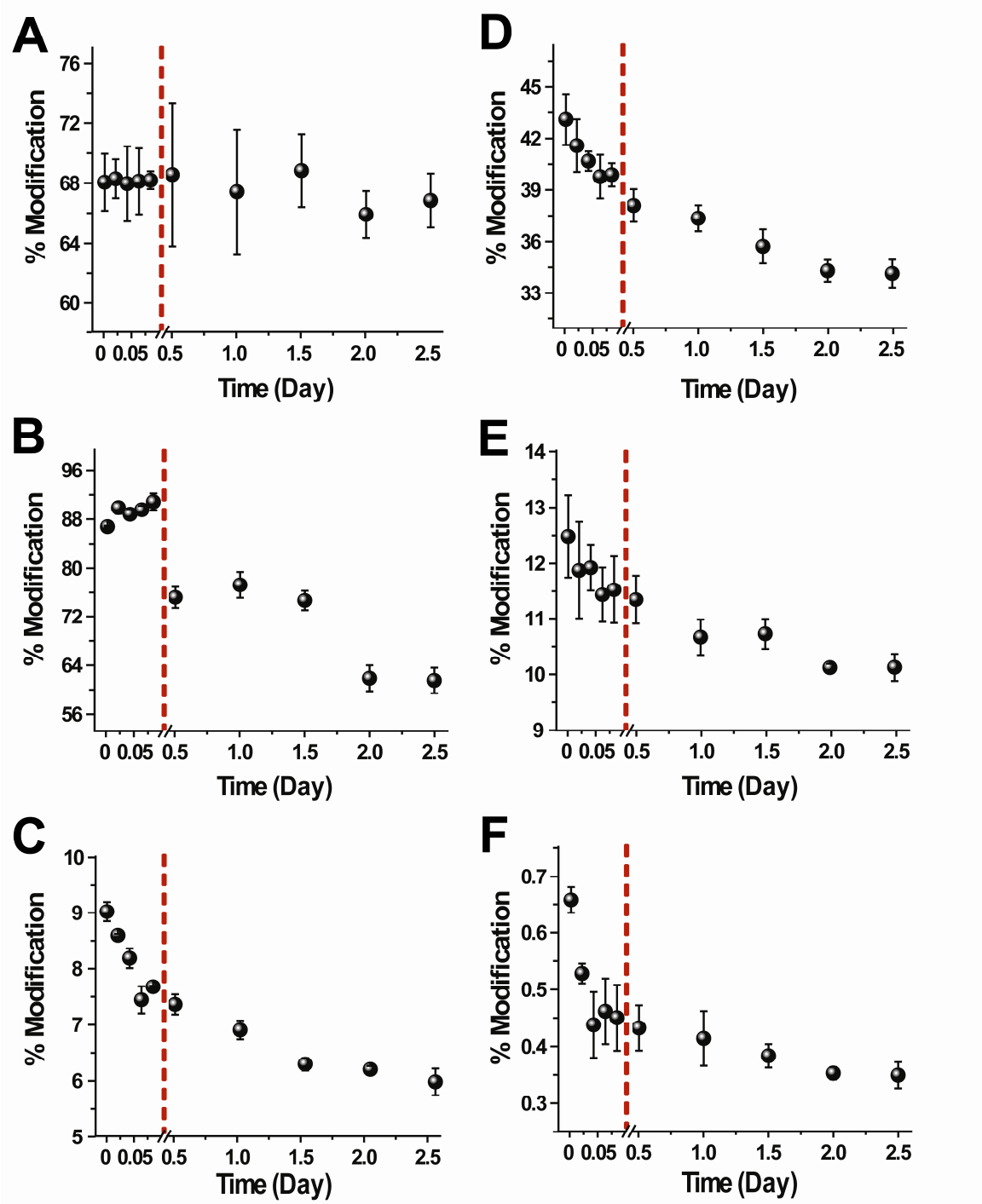
DEPC reacts readily with histidine residues but can also react with amine and hydroxyl groups at neutral pH. As was demonstrated previously, a 2.5-fold molar excess of DEPC is sufficient to label several amino acids in β2m while minimizing any DEPC-induced structural changes to the protein [10]. As was the case with the NHSA reactions, the reaction time was kept short (1 min) to avoid tetramer dissociation but also to minimize DEPC hydrolysis by water.

Proteolytic digestion of the protein and LC-MS/MS analyses indicate that DEPC reacts with the N-terminus, Thr4, Lys6, His13, Lys19, Tyr26, Ser28, His31, Ser33, Lys41, His51, Ser57/Lys58, Tyr63, Tyr67, Lys75, Ser88, and Lys94. These modifications are observed at all time points, but the level of modification for some residues changes over time (*vide infra*). In all cases but one (Ser57/Lys58), the specific amino acids that are modified could be determined unambiguously from the MS/MS data.

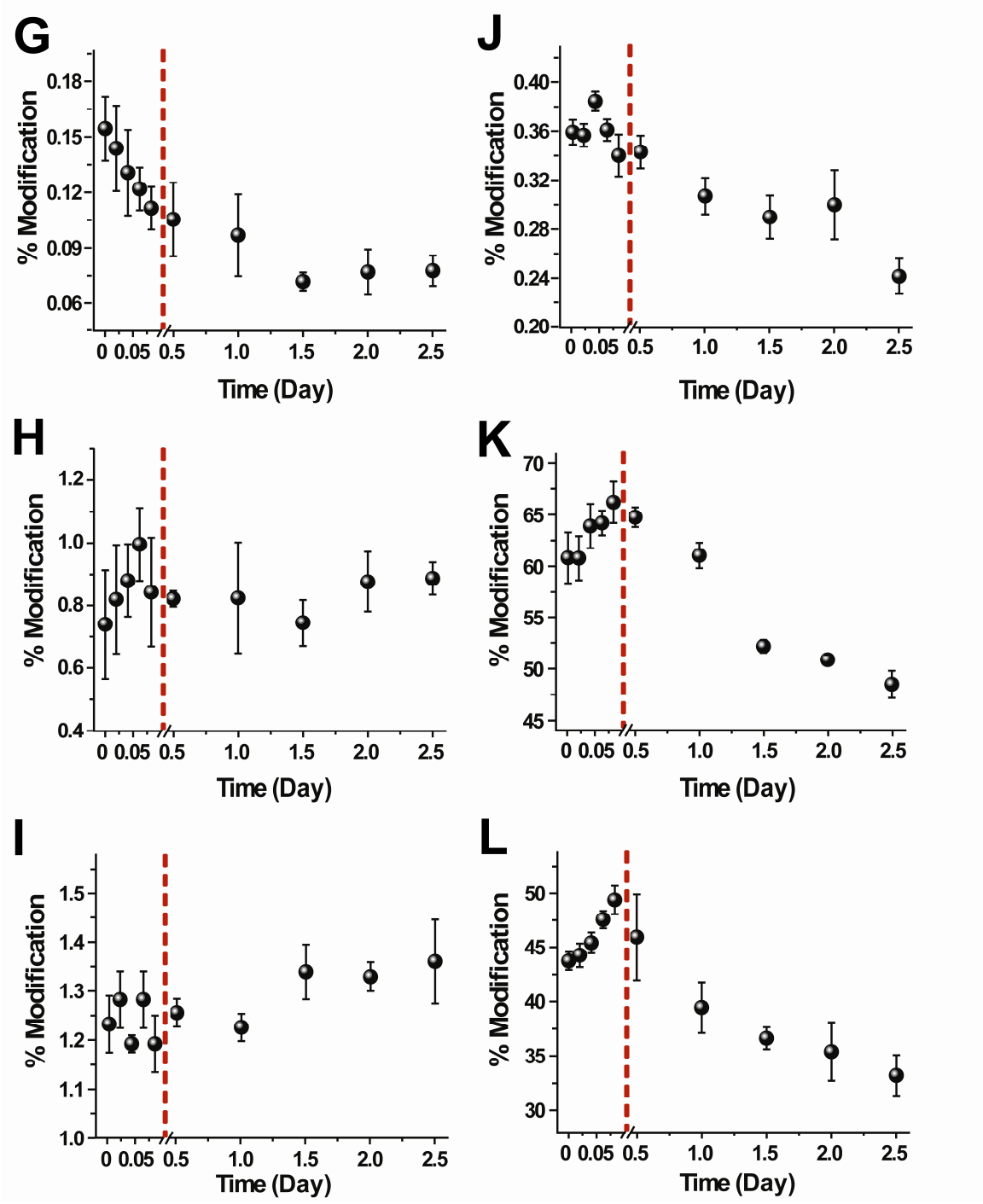
When the covalent labeling reactions with DEPC are performed at different time points after the tetramer begins to form, the modification levels of 13 of the 17 labeled residues change as the tetramer's concentration increases in solution (Figure 5.3). The labeling data reveal that only Lys75 (Figure 5.3O) shows an increase in modification extent upon tetramer formation. The level of modification of seven residues, Lys6, His13, Lys19, Tyr26, Ser28, Tyr63, and Tyr67, initially decrease within 2 hours after adding

Cu(II) and continue to decrease up to 2.5 days. Five residues, Thr4, Lys41, His51, Ser57/Lys58, and Lys94, increase within 2 hours after adding Cu(II) but then decrease as the tetramer concentration increases in solution. Four residues, the N-terminus, His31, Ser33, and Ser88, undergo no change in reactivity as the tetramer is formed. In comparing the DEPC and NHSA reactivity for those residues that react with both reagents, we find complete consistency for the N-terminus, Lys6, Lys41, Lys58, Lys75, and Lys94. Only the reactivity of Lys19 is slightly different.

In summary, the reactions with DEPC indicate that 13 residues undergo a notable change in reactivity as the tetramer is formed in solution. The modification extent of Lys75 increases whereas the reactivity of the other 12 amino acids decreases. The decrease in modification extents for Lys6, His13, Lys19, Tyr26, Ser28, Tyr63, and Tyr67 within 2 h after adding Cu(II) suggests that the continual drop in the extents of modification for these residues is due to formation of more dimer in solution rather than the formation of tetramer. Consequently, tetramer formation is then likely responsible for the reactivity changes of Thr4, Lys41, His51, Ser57/Lys58, Lys75, and Lys94.



This figure is continued on the next page.



This figure is continued on the next page.

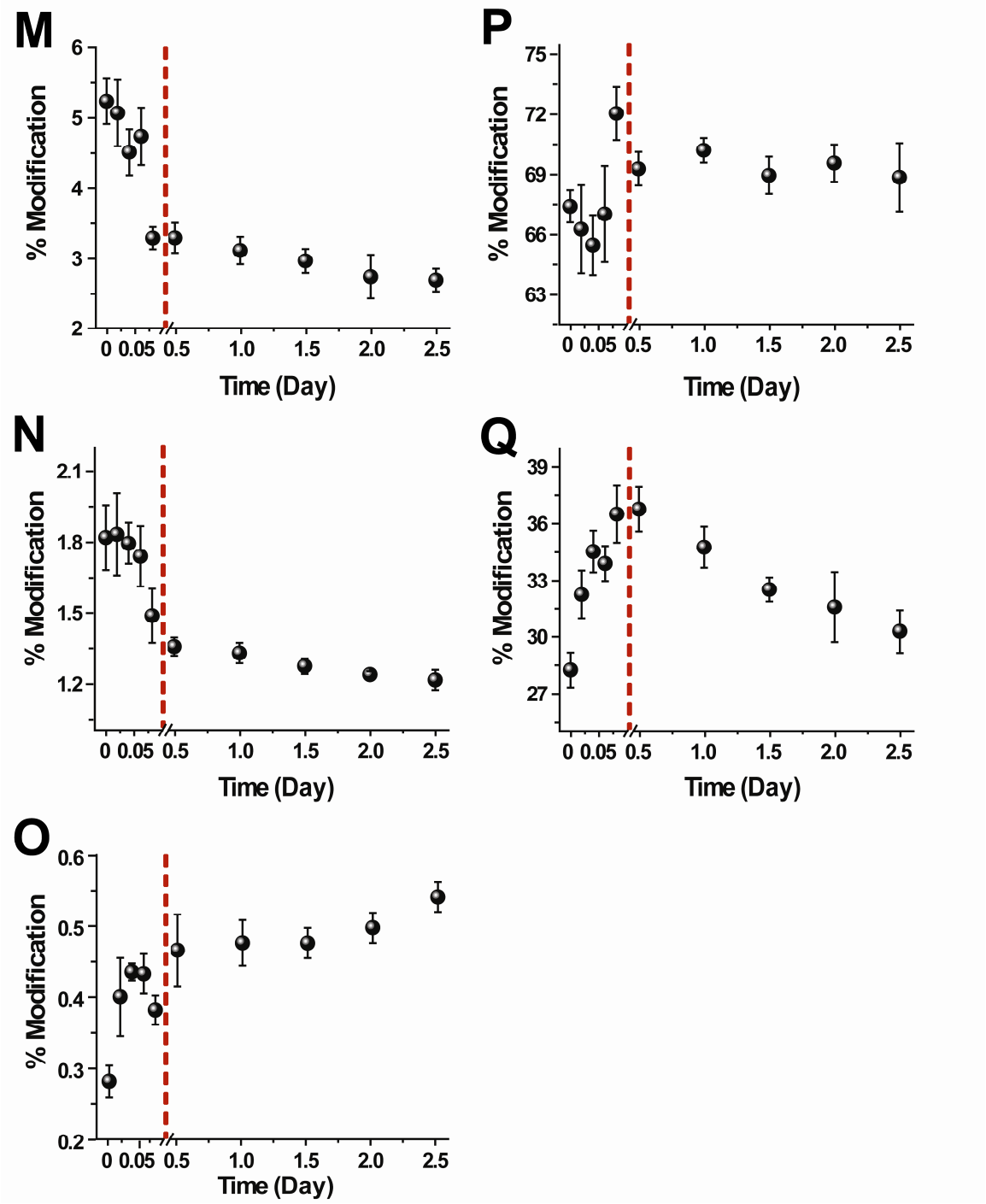


Figure 5.3 Extent of DEPC modification throughout the course of the tetramer formation. (A) N-terminus. (B) Th4. (C) Lys6. (D) His13. (E) Lys19. (F) Tyr26. (G) Ser28. (H) His31. (I). Ser33. (J) Lys41. (K) His51. (L) Ser57/Lys58. (M) Tyr63. (N) Tyr67. (O) Lys75. (P) Ser88. (Q) Lys94. The changes in modification during the dimer formation are shown (points before red line) as reference points.

5.3.3 Covalent Labeling with Butanedione

Unlike DEPC, which reacts with many nucleophilic groups, BD reacts specifically with arginine. The reaction is reversible at pH < 9, so higher reagent doses are necessary to improve the product yield. As determined previously, a reagent excess of 350 is suitable to avoid any modification-induced structural changes as well as to obtain readily detectable modifications [10]. The LC-MS data indicate that BD reacts with Arg3, Arg12, Arg45, and Arg97 but not Arg81. The reactions of BD with β2m reveal that only Arg3 and Arg12 undergo any notable changes in reactivity upon formation of the tetramer (Figure 5.4). The modification levels of Arg3 and Arg12 decrease as the tetramer's concentration increases in solution, but because the reactivities of these residues also decrease before 0.5 days their continual decreases in reactivity are likely due to formation of more dimer in solution rather than the formation of tetramer. The unchanged reactivity of Arg45 and Arg97 is curious since both of these residues increase in reactivity up to 2 h as the dimer's concentration increases in solution. It is likely that these residues slightly decrease in reactivity as the tetramer is formed, but this decrease is counterbalanced by the increase in reactivity as the dimer's concentration continues to increase in solution up to day 2.5. Overall, the BD reactivity suggests that Arg3 and Arg12 undergo no significant changes in reactivity as the tetramer is formed but Arg45 and Arg97 might.

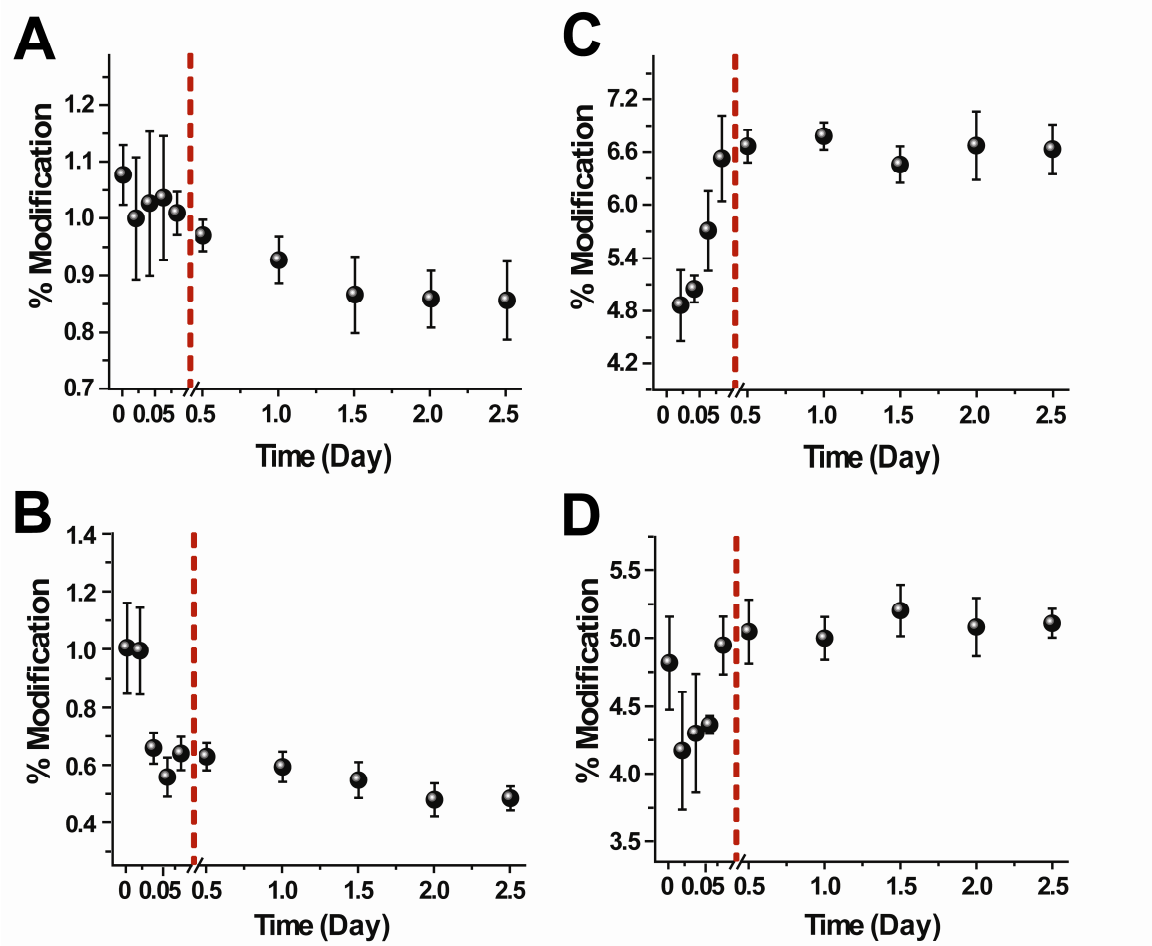


Figure 5.4 Extent of BD modification throughout the course of the tetramer formation. (A) Arg3. (B) Arg12. (C) Arg45. (D) Arg97. The changes in modification during the dimer formation are shown (points before red line) as reference points.

A summary of the reactivity changes for all of the modified amino acids monitored at t=2 h and t=2.5 days is shown in Table 5.1.

Table 5.1 Summary of the modification percentages for the modified amino acids at 2 hours (t=2 hours) after the addition of Cu(II) and 2.5 days (t=2.5 days) after the addition of Cu(II).

Residue	t = 2 hours	t = 2.5 days	% Change
NHSA			
N-terminus	60.7 ± 0.3	61 ± 1	0
Lys6	17.9 ± 0.1	14.2 ± 0.2	-
Lys19	7.2 ± 0.2	6.42 ± 0.09	-
Lys41	8.2 ± 0.8	6.1 ± 0.2	-
Lys58	19.0 ± 0.1	13.9 ± 0.5	-
Lys75	1.62 ± 0.08	2.33 ± 0.02	+
Asn83	1.46 ± 0.03	1.6 ± 0.1	+
Lys91	24.0 ± 0.4	21.0 ± 0.3	-
Lys94	4.4 ± 0.2	3.40 ± 0.08	-
DEPC			
N-terminus	68.1 ± 0.6	67 ± 2	0
Thr4	89.5 ± 0.6	61 ± 2	-
Lys6	7.64 ± 0.02	5.9 ± 0.2	-
His13	40.0 ± 0.7	34.2 ± 0.8	-
Lys19	11.5 ± 0.6	10.1 ± 0.2	-
Tyr26	0.45 ± 0.06	0.35 ± 0.02	-
Ser28	0.11 ± 0.01	0.078 ± 0.008	-
His31	0.8 ± 0.2	0.89 ± 0.05	0
Ser33	1.19 ± 0.06	1.36 ± 0.09	0
Lys41	0.34 ± 0.02	0.24 ± 0.01	-
His51	66 ± 2	49 ± 1	-
Ser57/Lys58	47 ± 2	33 ± 2	-
Tyr63	3.3 ± 0.2	2.7 ± 0.6	-

This table is continued on the next page.

Tyr67	1.5 ± 0.1	1.21 ± 0.04	-
Lys75	0.38 ± 0.02	0.54 ± 0.02	+
Ser88	72 ± 1	71 ± 2	0
Lys94	36 ± 2	28 ± 1	-
BD			
Arg3	1.01 ± 0.04	0.86 ± 0.07	-
Arg12	0.64 ± 0.06	0.48 ± 0.07	-
Arg45	6.5 ± 0.5	6.6 ± 0.3	0
Arg97	4.9 ± 0.2	5.1 ± 0.1	0

5.3.4 Pre-amyloid β2m Tetramer

The goal of this study is to obtain structural insight into the pre-amyloid tetramer that is formed by β2m after adding Cu(II) at near physiological conditions. Our group has previously shown that Cu(II)-induced β2m amyloid formation is preceded by the formation of dimers, tetramers, and hexamers [4]. No odd-numbered oligomers are observed, indicating that the oligomeric intermediates form via the assembly of dimers. In addition, our previous work suggests that the first tetramer species that is formed has Cu(II) bound to it, but as the amyloid fibril-forming reaction proceeds, Cu(II) is lost from this first tetramer before forming Cu(II)-free tetramers, hexamers, and eventually amyloid fibrils. In a recent study, we showed that the pre-amyloid dimer is formed by the interaction of anti-parallel ABED β-sheets from two monomers [10]. This dimer structure provides a starting point for gaining insight into the structure of the Cu(II)-bound tetramer.

Covalent modification with detection by MS was selected to probe the tetramer as this method has been effective in mapping protein surfaces, identifying ligand-binding sites, detecting ligand-induced conformational changes, and studying protein-protein complexes, including the β2m dimer [10,19]. As described in our previous studies [10], care was taken to select optimal reaction times and stoichiometries for each covalent labeling reagent to ensure that the amount of label added and the resulting modification do not disrupt the protein structure. A total of 24 residues in β2m are probed by NHSA, DEPC, and BD, and these residues are widely distributed along the polypeptide chain and on the surface of the protein, representing about 30% of the surface amino acids (see monomer in Figure 5.1A). Changes in the reactivity of amino acid side chains reflect changes in their solvent accessibility. Hence, patterns of modification during the course of the amyloid fibril-formation reaction permit us to build a low-resolution map of the protein and identify residues involved in mediating tetramer formation. Having baseline labeling data for the monomer and dimer allow us to readily identify those residues involved in tetramer formation, even though monomer, dimer, and tetramer are present simultaneously. In addition, being able to perform these reactions on wild-type β2m under amyloid fibril-forming conditions makes the resulting structural information relevant for understanding native β2m amyloid assembly.

The amino acid side chains probed by the covalent labels can be grouped in three categories, namely, residues with (1) unchanged reactivity, (2) decreased reactivity, and (3) increased reactivity from 2 h after adding Cu(II) to 2.5 days when the first form of the tetramer is present. The covalent labeling experiments reveal that six of the 24 modified amino acids, the N-terminus, His31, Ser33, Arg45, Ser88, and Arg97, maintain the same

reactivity with their respective labeling reagents at time points from 2 h to 2.5 days. Four of these residues, the N-terminus, His31, Ser33, and Ser88 have the same reactivity in the dimer and tetramer, indicating that these residues retain similar microenvironments in both these oligomers. Therefore, these residues are not likely part of the tetramer interface. The unchanged reactivities of the N-terminus and His31 are consistent with these residues remaining Cu(II) binding sites in the monomer, dimer, and tetramer [20]. The unchanged reactivity of Ser33 is also consistent with steric hindrance caused by nearby Cu(II) binding in the monomer and oligomers. Others have reported that Cu(II) binding induces a *cis*-to-*trans* backbone isomerization of Pro32. In the crystal structures of the P32A mutant dimer and the dimeric unit in the H13F mutant hexamer, which both show this cis-trans isomerization, Ser33 has low accessibility to solvent [8,9]. Persistence of this structural feature in the tetramer due to Cu(II) binding would maintain the microenvironment around Ser33 in such a way that its reactivity does not change. The unchanged reactivity of Ser88 suggests that it remains solvent accessible in the monomer, dimer, and tetramer. The reactivity of Arg45 and Arg97 is essentially different than the N-terminus, His31, Ser33, and Ser88. These residues increase in reactivity as the dimer is formed but then their reactivities level off as the tetramer is formed. This behavior suggests that they decrease in reactivity as the tetramer is formed, thereby counterbalancing their increased reactivity as the dimer continues to populate the solution. Thus, we consider Arg45 and Arg97 as residues that decrease in reactivity as the tetramer is formed.

Including Arg45 and Arg97, 18 residues decrease in reactivity as the tetramer's concentration increases. Nine of these residues, Arg3, Lys6, Arg12, His13, Lys19, Tyr26, Ser28, Tyr63, and Tyr67, were found previously to decrease in reactivity as the dimer is formed, and their continued decrease in reactivity is consistent with the dimer's continued increase in concentration up to 2.5 days. The remaining nine residues, Thr4 on the A strand, Lys41 on the C strand, Arg45 on the CD loop, His51 on the D strand, Ser57/Lys58 on the DE loop, Lys91, Lys94, and Arg97 on the G strand, only decrease in reactivity once the tetramer is formed. Because solvent accessibility is the main factor that affects the reactivity of the amino acid side chains, decreases in reactivity suggests that these latter nine amino acids may mediate interactions in the tetramer.

Even though no crystal structure of a wild-type or mutant β2m tetramer is available to which we can compare our covalent labeling data, several lines of evidence help us arrive at a model for the pre-amyloid tetramer of the wild-type protein. (1) Previous reports have shown that the β2m oligomers are native-like in structure [21], indicating that the tetramer will have many of the structural features present in the wild-type monomer for which several crystal structures exist [22,23]. (2) Previous covalent labeling experiments from our group determined the structure of the wild-type β2m dimer, which has essentially the same structural features as the dimer unit formed by the B and C chains of the H13F mutant hexamer of β2m [10]. This dimer structure (see Figure 5.1) serves as a convenient starting point for constructing a model of the tetramer since the tetramer very likely arises from the assembly of two dimers. (3) The H13F mutant hexamer contains another oligomeric interface (chains C and D) that could represent the interface of the wild-type tetramer and so our labeling data can be directly

compared to the interface of these two chains. (4) The P32A mutant dimer, which was previously ruled out as a possible dimer interface for wild-type β2m, could represent the interface of the wild-type tetramer and our labeling data can be compared to this structure too.

In comparing our labeling data with the oligomeric interface of chains C and D of the H13F hexamer (Figure 5.5), we find that only three of the nine residues that decrease in reactivity are found at the CD interface. The interface of these two chains should result in reduced solvent accessibility for His51, Ser57, and Lys58, and we observe experimentally that these residues decrease in reactivity upon tetramer formation. In contrast, the other six residues (i.e. Thr4, Lys41, Arg45, Lys91, Lys94, and Arg97) that decrease in reactivity and thus are likely at the tetramer interface, are all solvent accessible when considering the CD interface of the H13F hexamer (Figure 5.5C). This collective observation alone strongly suggests that the interface of the C and D chains of the H13F mutant does not reflect the tetramer interface in wild-type β2m. Indeed, no interface in the H13F hexamer is consistent with our covalent labeling data for the wild-type tetramer. This conclusion is not too troublesome, though, as the H13F hexameric mutant cannot form amyloid fibrils, whereas the wild-type protein can form tetramers, hexamers and eventually amyloid fibrils. Evidently, the wild-type protein forms a tetramer species capable of continuing on to form amyloid fibrils, whereas the H13F mutant cannot form such an interface and therefore does not progress to form amyloid fibrils.

Since studies have shown that the H13F mutant has a very similar structure, stability, and Cu(II) binding constant compared to wild-type β2m [24], the inability of H13F mutant to form amyloid fibrils indicates that His13 is important in the formation of pre-amyloid oligomers of wild-type β2m. The exact role that His13 plays is not clear, but the presence of a Cu-bound and a Cu-free tetramer [4] along with the metal-binding capabilities of histidine residues might provide a clue. It is possible that transiently binding of His13 by Cu(II) may provide the necessary stabilization to allow release of Cu(II) from the N-terminal region, thus enabling the formation of the Cu(II)-free tetramer that eventually proceeds to form the Cu(II)-free amyloid fibrils. Replacement of His by Phe in the H13F mutant may hinder the formation of the Cu(II)-free tetramer, which is necessary for eventual fibril formation.

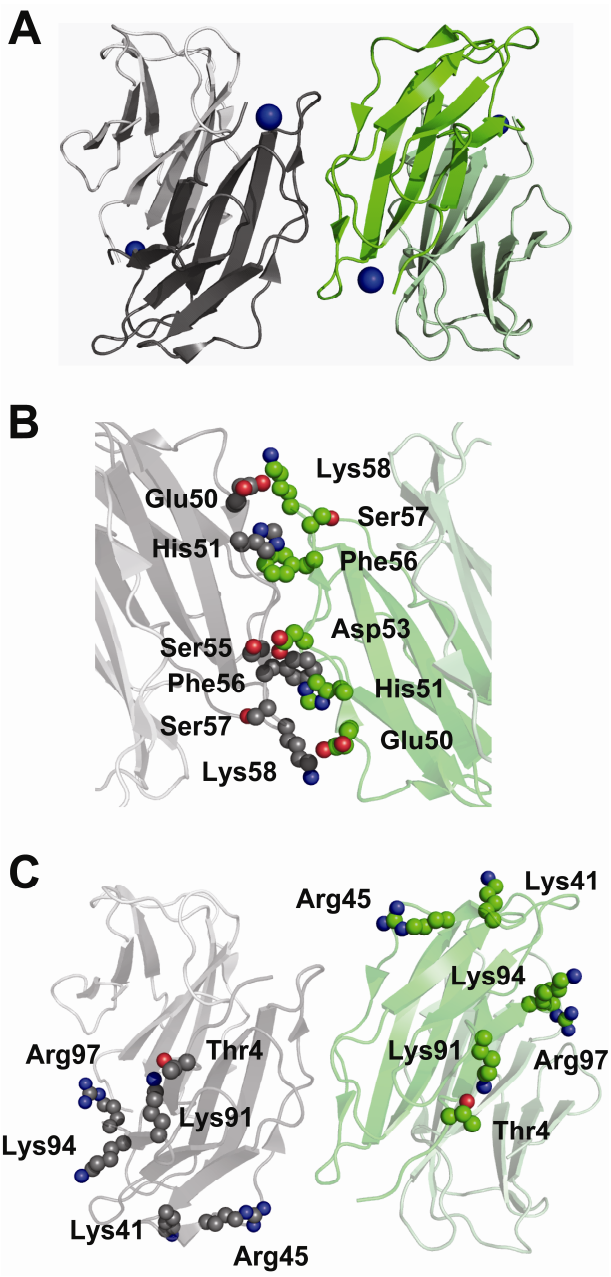


Figure 5.5 Possible tetramer formed by two dimer units with anti-parallel ABED interface as seen in the H13F hexamer [9]. The side chains involved in the interface are shown as spheres. (A) Possible tetramer formed by D strands and BC loops from two dimer units in the crystal structure of H13F hexamer. One dimer unit (chains BC) is shown in gray and the other (chains DE) is shown in green. Chains C and D, which form tetramer interface, are shown in dark gray and dark green, respectively. (B) Interaction of adjacent D strands. Inter-strand interactions of the side chains of Glu50 and His51 of one monomer with the side chain of Lys58 and the backbone of Phe56, respectively, of another monomer are shown. These interactions make Ser57 less exposed to solvent. (C) Six of the amino acids that decrease in reactivity upon tetramer formation, namely, Thr4, Lys41, Arg45, Lys91, Lys94, and Arg97, are solvent accessible in the H13F structure.

The P32A mutant also forms an oligomeric interface with which we can compare our labeling data (Figure 5.6). Like the interactions seen in the CD chains of the H13F hexamer, the interaction of the anti-parallel D strands in P32A span primarily the amino acids Glu50-Lys58 (Figure 5.6B). These interfaces predict reduced solvent accessibility of His51 on the D strand and Ser57 and Lys58 on the DE loop. These interactions protect residues His51, Ser57, and Lys58 from solvent, which possibly explains why we see a decrease in the reactivity of these residues. Like the H13F mutant, though, Thr4, Lys41, Arg45, Lys91, Lys94, and Arg97 are solvent accessible in the P32A dimer, which is not consistent with the observed decrease in their reactivity upon tetramer formation. This result indicates that an interaction involving D strands from two protein sub-units does not represent the tetramer interface in wild-type β2m.

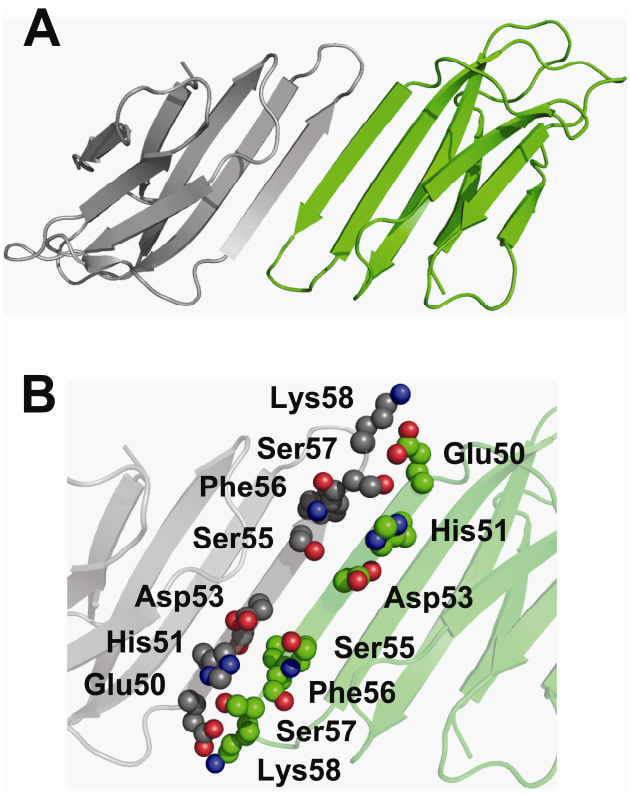


Figure 5.6 Interface formed by D-D strand interactions in P32A. (A) Crystallographic dimer formed by P32A. (B) Amino acids Glu50-Lys58 on the D strands are shown as spheres. Inter-strand interactions of the side chains of Glu50 and His51 of one monomer with the side chain of Lys58 and the backbone of Phe56, respectively, of another monomer are shown.

In arriving at a viable model for the tetramer, it is helpful to map the amino acids that decrease in reactivity upon tetramer formation. In doing so, it is clear that the residues that decrease in reactivity are localized on the edge strands A, D, C and G strands of β2m (Figure 5.7). This observation suggests a tetramer interface that involves an anti-parallel arrangement of D and G strands from two dimers (Figure 5.8). In this tetramer model, two D strands from one dimer unit and two G strands from an adjacent dimer form the interface (Figure 5.8A and 5.8B). More specifically, a D strand from one

monomer in a dimer and a D strand from the other monomer in the same dimer interact with the G strands from each of the monomers the other dimer. Analysis of such a model, which was computationally constructed and energy minimized as described in the experimental section, reveals that seven out of the nine residues that drop in reactivity as the tetramer is formed are found at the interface of the D and G strands of adjacent dimers. These seven residues are Thr4, His51, Ser57/Lys58, Lys91, Lys94, and Arg97 (Figure 5.8C). The two amino acids that increased in reactivity upon tetramer formation, Lys75 and Asn83, are both solvent accessible in the tetramer model. In this model, His51, Lys58, Lys91, Lys94, and Arg97 are involved in mediating the interactions at the tetramer interface, whereas Ser57 is protected from solvent by these interactions and Thr4 is enclosed by the D and G strands. Glu50 and His51 form salt bridges with Arg97 and Asp96, respectively. Such electrostatic interactions can substantially enhance the stability of protein-protein complexes [25,26], and the involvement of His51 in one of these interactions is consistent with a study by Blaho and Miranker in which they found from mutagenesis studies that His51 is a critical residue in mediating the interface [27]. Asp53 is in close proximity to both Lys91 and Lys94, indicating that this residue can form a complex salt bridge with both of these lysine residues. Such complex salt bridges at protein-protein interfaces contribute significantly to protein complex stability [28-30]. In addition, Lys58 is in close enough proximity to hydrogen bond with Gln89. Moreover, several hydrophobic residues are buried in this model structure, including Pro32, Leu54, Phe56, Ile92, and Val93 (residues not shown). In this model, Arg45 decreases in reactivity but is not buried at the interface. In this case, the orientation of the D strand is altered such that Arg45 can form a salt bridge with Glu47 (Figure 5.8D), which could

explain its decreased reactivity with BD. These interactions might also alter the orientation of the D strand such that Lys41 becomes protected from solvent, and thus decrease in reactivity upon tetramer formation.

A final important feature of our proposed model for the pre-amyloid tetramer is the likely ability of this tetramer to further propagate into higher order oligomers. In our previous studies, we found that the tetramer can progress to form hexamers. As shown in Figure 5.8, interaction of D strands of one dimer and the G strands of a second dimer to form the tetramer leaves the G strands of the first dimer and the D strands of the second dimer available for further interaction with another dimer unit. Such unfulfilled interactions would presumably allow formation of a hexamer via another D-G interaction.

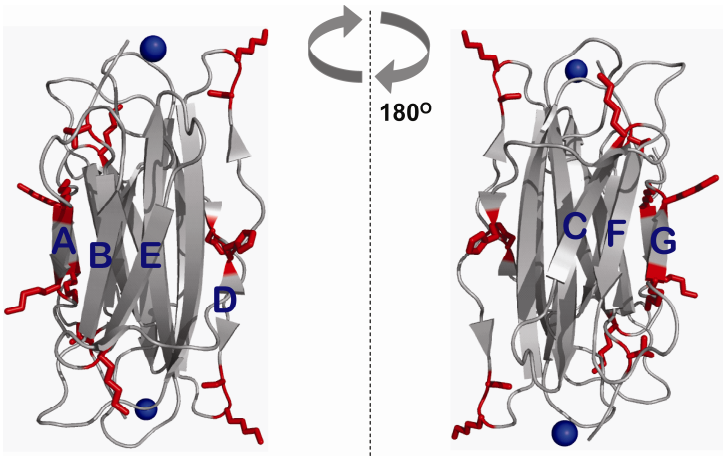


Figure 5.7 The probed amino acids that decrease in modification upon tetramer formation (shown as red sticks) are located on the edge strands, A, C, D and G.

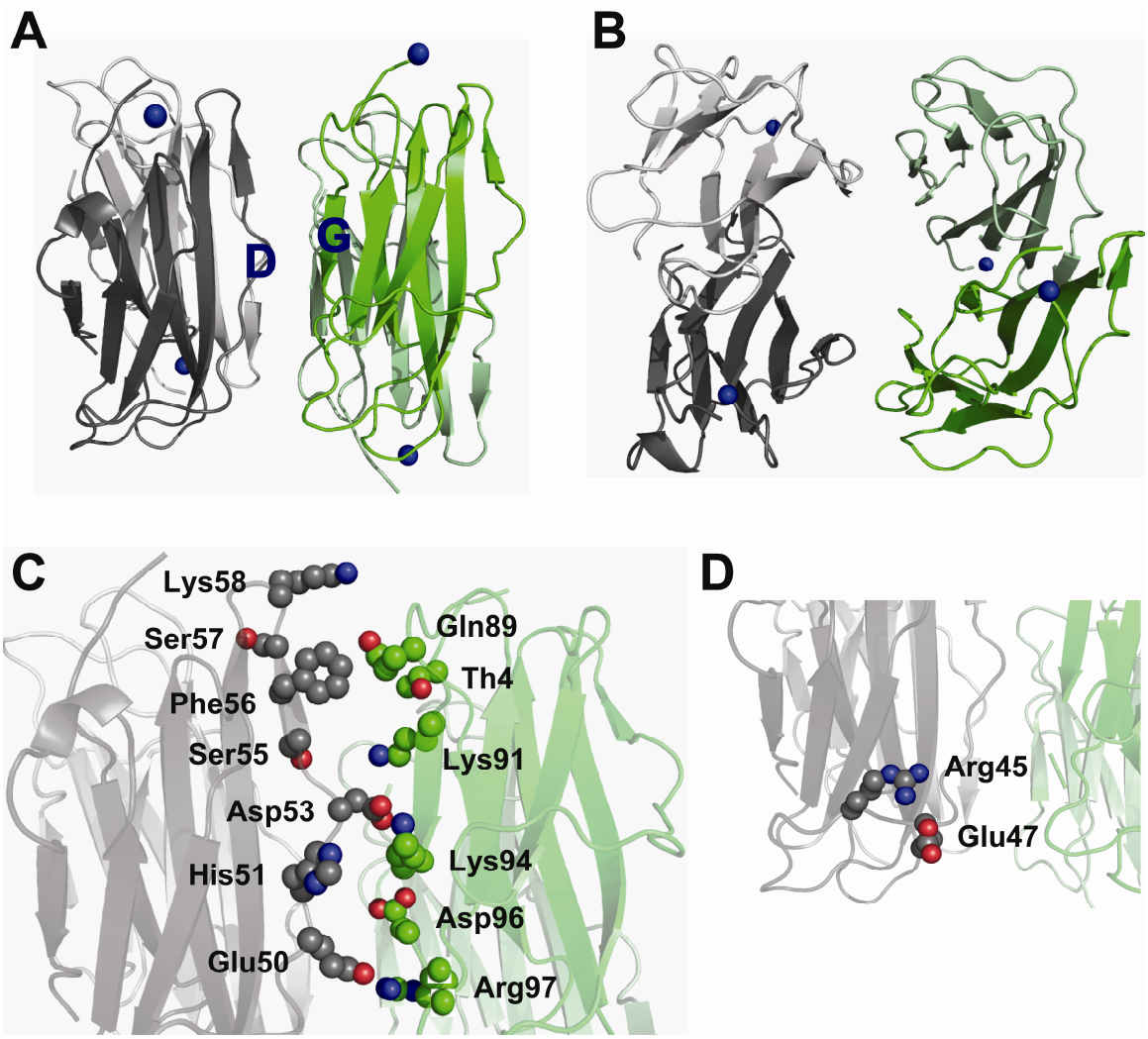


Figure 5.8 Proposed model structure for the wild-type β2m tetramer. (A) Tetramer formed by the interaction of two D strands from one dimer unit (dark and light gray) and two G strands of another dimer unit (dark and light). (B) Top-view of the tetramer model. (C) Amino acids Thr4 on the N-terminal strand, Glu50-Lys58 on the D strand, and Gln89-Arg97 on the G strand are shown as spheres. Several interactions are shown: salt bridge between Glu50 and Arg97, salt bridge between His51 and Asp96, complex salt bridge between Asp53 and Lys94 and Lys91, and H-bonding between Gln89 and Lys58. (D) These interactions alter the orientation of the D strand such that Arg45 can form a salt bridge with Glu47.

5.4 Conclusions

Using a combination of selective covalent labeling and mass spectrometric detection, we identify structural features of the pre-amyloid tetramer of β2m, which is generated by incubation with Cu(II) under physiologically relevant conditions. The covalent labeling data allows us to develop a model for the tetramer in which the interface is mediated by interactions between D strands of one dimer unit and the G strands of another dimer unit. Our model, therefore, differs from previous suggestions that the tetramer is mediated by interactions between D strands on separate dimer units. Covalent labeling data that covers about 30% of the amino acids that are at the interface of the proposed tetramer structure provide strong support for our model. Future studies will further test this model, and more structural information about this pre-amyloid tetramer might be useful for the development of therapeutics against DRA. More broadly speaking, covalent labeling along with MS detection should have applicability to other amyloid-forming systems, especially those in which mixtures of pre-amyloid oligomers are formed. This approach could then take on greater importance since in several amyloid diseases the pre-amyloid oligomers may represent the species responsible for cellular toxicity rather than the amyloid fibrils themselves [31,32].

5.5 References

1. Ferrao-Gonzales, A.D., Robbs, B.K., Moreau, V.H., Ferreira, A., Juliano, L., Valente, A.P., Almeida, F.C.L., Silva, J.L., and Foguel, D. (2005) Controlling β-amyloid oligomerization by the use of naphthalene sulfonates—trapping low molecular weight oligomeric species. *J. Biol. Chem. 280*, 34747–34754.

2. Friedhoff, P., von Bergen, M., Mandelkow, E.-M., Davies, P., and Mandelkow, E. (1998) A nucleated assembly mechanism of Alzheimer paired helical filaments. *Proc. Natl. Acad. Sci USA. 95*, 15712–15717.

3. Walsh, D.M., Lomakin, A., Benedek, G.B., Condron, M.M., and Teplow, D.B. (1997) Amyloid b-protein fibrillogenesis—detection of a protofibrillar intermediate. *J. Biol. Chem.* 272, 22364–22372.

4. Antwi, K., Mahar, M., Srikanth, R., Olbris, M.R., Tyson, J.F., and Vachet, R.W. (2008) Cu(II) organizes beta-2-microglobulin oligomers but is released upon amyloid formation. *Protein Sci. 17*, 748-759.

5. Eakin, C.M., Knight, J.D., Morgan, C.J., Gelfand, M.A., and Miranker, A.D. (2002) Formation of a copper specific binding site in non-native states of β-2-microglobulin, *Biochemistry 41*, 10646-10656.

6. Bucciantini, M., Giannoni, E., Fabrizio, C., Baroni, F., Formigli, L., Zurdo, J., Taddei, N., Ramponi, G., Dobson, C.M., and Stefani, M. (2002) Inherent toxicity of aggregates implies a common mechanism for protein misfolding diseases. *Nature 416*, 507-511.

7. Mendes Sousa, M., Cardoso, I., Fernandes, R., Guimaraes, A., and Saraiva, M.J. (2001) Deposition of transthyretin in early stages of familial amyloidotic polyneuropathy. Evidence for toxicity of nonfibrillar aggregates. *Am. J. Pathol. 159*, 1993-2000.

8. Calabrese, M.F., and Miranker, A.D. (2007) Formation of a stable oligomer of β-2-microglobulin requires only a transient encounter with Cu(II). *J. Mol. Biol. 367*, 1-7.

9. Eakin, C.M., Berman, A.J., and Miranker, A.D. (2006) A native to amyloidogenic transition regulated by a backbone trigger. *Nature Struct. Mol. Biol. 13*, 202-208.

10. Mendoza, V.L., Antwi, K., Baron-Rodriguez, M.A., Blanco, C., and Vachet, R.W. (2009) Structure of the preamyloid dimer of beta-2-microglobulin from covalent labeling and mass spectrometry. *Biochemistry 49*, 1522-1532.

11. Cornell, W., Cieplak, P., Bayly, C., Gould, I., Merz, K., Ferguson, D., Spellmeyer, D., Fox, T., Caldwell, J., and Kollma, P. (1995) A second generation force field for the simulation of proteins, nucleic acids, and organic molecules. *J. Am. Chem. Soc. 117*, 5179-5197.

12. Sorin, E.J. and Pande, V.S. (2005) Exploring the helix-coil transition via all-atom equilibrium ensemble simulations. *Biophys. J. 88*, 2472-2493.

13. van der Spoel, D., Lindahl, E., Hess, B., Groenhof, G., Mark, A., and Berendsen, H. (2005) Gromacs: Fast, flexible, and free. *J. Comp. Chem. 26*, 1701-1719.

14. Lindahl, E., Hess, B., and van der Spoel, D. (2001) Gromacs 3.0: A package for molecular simulation and trajectory analysis. *J. Mol. Mod. 7*, 306-317.

15. Berendsen, H., van der Spoel, D., and van Drunen, R. (1995) Gromacs: A message-passing parallel molecular dynamics implementation. *Comp. Phys. Comm. 91*, 43-56.

16. Verdone, G., Corazza, A., Viglino, P., Pettirossi, F., Giorgetti, S., Mangione, P., Andreola, A., Stoppini, M., Bellotti, V., and Esposito, G. (2002) The solution structure of human β2-microglobulin reveals the prodromes of its amyloid transition. *Protein Sci. 11*, 487-499.

17. Mahoney, M. Jorgensen, A. (2000) A five-site model for liquid water and the reproduction of the density anomaly by rigid, nonpolarizable potential functions. *J. Chem. Phys. 112*, 8911-8922.

18. Okon, M., Bray, P., and Vucelic, D. (1992) H^1-NMR assignments and secondary structure of human beta-2-microglobulin in solution. *Biochemistry 31*, 8906−8915.

19. Mendoza, V. L., and Vachet, R. W. (2008) Protein surface mapping using diethylpyrocarbonate with mass spectrometric detection. *Anal. Chem. 80*, 2895-2904.

20. Srikanth, R., Mendoza, V. L., Bridgewater, J. D., Zhang, G., and Vachet, R. W. (2009) Copper binding to β-2-microglobulin and its pre-amyloid oligomers. *Biochemistry 48*, 9871-9881.

21. Eakin, C. M., Attenello, F. J., Morgan, C.J., and Miranker, A.D. (2004) Oligomeric assembly of native-like precursors precedes amyloid formation by β-2-microglobulin. *Biochemistry 43*, 7808-7815.

22. Trinh, C. H.; Smith, D. P.; Kalverda, A. P.; Phillips, S. E. V. and Radford, S. E. (2002) Crystal structure of monomeric human β-2-microglobulin reveals clues to its amyloidogenic properties. *Proc. Natl. Acad. Sci. USA 99*, 9771-9776.

23. Kihara, M., Chatani, E., Iwata, K., Yamamoto, K., Matsuura, T., Nakagawa, A., Naiki, H., and Goto, Y. (2006) Conformation of amyloid fibrils of β2-microglobulin probed by tryptophan mutagenesis. *J. Biol. Chem. 281*, 31061-31069

24. Eakin, C.M., Knight, J.D., Morgan, C.J., Gelfand, M.A., and Miranker, A.D. (2002) Formation of a copper specific binding site in non-native states of β-2-microglobulin. *Biochemistry 41*, 10646-10656.

25. Kundrotas, P. J., and Alexov, E. (2006) Electrostatic Properties of Protein-Protein Complexes. *Biophys. J. 91*, 1724–1736.

26. Dong, F., and Zhou, H.-X. (2006) Electrostatic Contribution to the Binding Stability of Protein–Protein Complexes. *Proteins 65*, 87–102.

27. Blaho, D. V., and Miranker, A. D. (2009) Delineating the conformational elements responsible for Cu2+-induced oligomerization of beta-2 microglobulin. *Biochemistry 48*, 6610-6617.

28. Musafia, B., Buchner, V., and Arad, D. (1995) Complex salt bridges in proteins: Statistical analysis of structure and function. *J Mol. Biol. 254*, 761-770.

29. Risal, D., Gourinath, S., Himmel, D. M., Szent-Gyorgyi, A. G., and Cohen, C. (2004) Myosin subfragment 1 structures reveal a partially bound nucleotide and a complex salt bridge that helps couple nucleotide and actin binding. *Proc. Natl. Acad. Sci. USA 101*, 8930-8935.

30. Xu, G. Z., Liu, R. T., Zak, O., Aisen, P., and Chance, M. R. (2005) Structural allostery and binding of the transferrin·receptor complex. *Mol. Cell. Proteomics 4*, 1959-1967.

31. Bucciantini, M., Giannoni, E., Fabrizio, C., Formigli, L., Zurdo, J., Taddei, N., Ramponi, G., Dobson, C.M., and Stefani, M. (2002) Inherent toxicity of aggregates implies a common mechanism for protein misfolding disease. *Nature 416*, 507-511.

32. Mendes Sousa, M., Cardoso, I., Fernandes, R., Guimaraes, A., and Saraiva, M.J. (2001) Deposition of transthyretin in early stages of familial amyloidotic polyneuropathy. Evidence for toxicity of nonfibrillar aggregates. *Am. J. Pathol. 159*, 1993-2000.

CHAPTER 6

DELINEATING THE METAL-SPECIFIC STRUCTURAL CHANGES IN β-2-MICROGLOBULIN AND THEIR EFFECT ON ITS AGGREGATION

6.1 Introduction

Prior work shows that endogenous metals other than Cu, such as Fe and Zn, are involved in the pathogenesis of a variety of aggregation and amyloid diseases. Reports have shown that Cu(II) and Zn(II) modulate the aggregation of the prion peptide PrP106-126 from Creutzfeldt-Jakob disease [1]. It has been shown that Aβ aggregation is greatly influenced by Zn(II) at pH 7.4. Cu(II) and Fe(III), which have much lower affinities for the protein, can induce limited Aβ aggregation, but their effect can be further exaggerated by slightly acidic conditions [2,3]. Studies have shown that α-synuclein from Parkinson's is able to interact with Zn(II), Cu(II), and Fe(III) and that these interactions lead to protein aggregation and crosslinking [4,5]. The amyloidogenic immunoglobulin light chain (LC) or SMA forms granular aggregates upon exposure to Cu leading to LC deposition disease [6]. This previous work demonstrates that transition metals can influence protein amyloid formation and raises the question about the effect of other transition metals on β2m aggregation.

As shown in the previous chapters, interaction of β2m with Cu(II) results in destabilization, oligomerization and subsequent amyloid formation. In this chapter, the effect of interactions with other first-row transition metals, specifically Zn(II) and Ni(II), on β2m will be investigated. Previous studies show that Zn(II) and Ni(II) also bind to β2m but with much lower affinities (K_d^{Cu} = 2.7 μM, K_d^{Ni} = 400 μM and K_d^{Zn} = 1500 μM)

[7]. Importantly, the binding of Ni(II) and Zn(II) to β2m does not lead to amyloid fibril formation, but Zn(II) does induce the formation of oligomers [8]. In addition to native state binding, urea titration experiments shows that the destabilization of β2m is specific for Cu(II) over Zn(II) and Ni(II) [7,9]. Taken as a whole, these previous reports indicate that Cu(II) exerts a unique effect by stimulating β2m amyloid formation. It is therefore interesting to compare the differences that each metal (Cu(II), Ni(II), and Zn(II)) exerts on the structure of the β2m monomer as a way to obtain deeper insight into the specific structural changes that are necessary for Cu(II) to induce β2m amyloid formation.

6.2 Experimental Procedure

6.2.1 Materials

Human β2m was obtained from Fitzgerald Industries International, Inc. (Concord, MA). Diethylpyrocarbonate (DEPC), 2,3-butanedione (BD), sulfo-*N*-hydroxysuccinimide acetate (NHSA), copper(II) sulfate ($CuSO_4$) pentahydrate, zinc sulfate ($ZnSO_4$) heptahydrate, nickel sulfate ($NiSO_4$) heptahydrate, 3-morpholinopropanesulfonic acid (MOPS), imidazole, dithiothreitol (DTT), potassium acetate, arginine, carbonic anhydrase, equine skeletal muscle myoglobin, and bovine serum albumin were purchased from Sigma-Aldrich (St. Louis, MO). Tris(hydroxymethyl)-aminomethane (Tris) was purchased from EM Science (Gladstone, NJ). Urea was purchased from Mallinckrodt Chemicals (Phillipsburg, NJ). Trypsin was from Promega (Madison, WI), and chymotrypsin was purchased from Roche Diagnostics (Indianapolis, IN). Centricon molecular weight cutoff filters were obtained from Millipore (Burlington, MA). Deionized water was prepared from a Millipore Simplicity 185 water purification system.

6.2.2 Formation of β2m Oligomers

Oligomers were formed by incubation of 100 μM β2m in 200 mM potassium acetate, 500 mM urea, and 25 mM MOPS (pH 7.4) with the divalent metal salt of interest at 37 °C. The concentrations of the metal salts used were: 200 μM $CuSO_4$, 400 μM $ZnSO_4$, and 16 mM $NiSO_4$. All components were equilibrated at 37 °C prior to metal salt addition, and immediately returned to 37 °C after mixing. The covalent labels were added to aliquots of incubated β2m taken at 2 minutes after addition of the metal salt. Five μL aliquots of the incubated β2m were taken at several time points after addition of the metal salt for analysis by size-exclusion chromatography (SEC).

6.2.3 Covalent Labeling Reactions

The covalent labeling reactions were done in a manner identical to that described in Section 3.2.3-3.2.5.

6.2.4 Proteolytic Digestion

The proteolytic digestions were done in a manner identical to that described in Section 3.2.6.

6.2.5 Instrumentation

The determination of the amount of modification of the whole protein and proteolytic fragments were performed the same way as described in Section 3.2.7. For online SEC-MS analyses, a Bruker Amazon quadrupole ion trap mass spectrometer equipped with a standard ESI source was used. The ESI source was operated at a spray

voltage of 4.5 kV, and the capillary temperature was set at 220°C. The voltages for the transfer optics between the ESI source and the ion trap were optimized for maximum signal.

6.2.6 Amino Acid Modification Percentage

The amino acid modification percentages were determined in a manner identical to that described in Section 3.2.8.

6.2.7 Circular Dichroism (CD)

CD measurements were made at 37 ^{o}C using a Jasco-710 spectropolarimeter. The metal salts (same concentrations used as in section 6.2.2) were added to a solution containing 20 μM β2m, 500 mM Urea, 25 mM MOPS and 150 mM potassium acetate. Spectra were acquired with a sample cell having a 0.1 cm path length and at a scan resolution of 0.2 nm, a scan rate of 50 nm/min, and a response time of 4 s.

6.2.8 Thioflavin T Fluorescence

Fluorescence experiments to monitor the rate of β2m oligomer formation were performed using a QuantaMaster 4 SE spectrofluorometer (Photon Technology International, Lawrenceville, NJ). A solution containing 100 μM β2m, 80 μM thioflavin T (ThT), 500 mM Urea, 25 mM MOPS and 150 mM potassium acetate was initially equilibrated at 37 ^{o}C, and after the addition of the divalent metal salt, the fluorescence shift of ThT was measured. Measurements were taken at excitation and emission wavelengths of 437 and 483 nm, respectively.

6.2.9 Size Exclusion Chromatography

To determine any β2m oligomers formed upon addition of the divalent metals, the incubated solutions were separated using a Superdex 75 PC 3.2/30 column (GE Healthcare Biosciences, Piscataway, NJ) installed on an Agilent HP1100 series high-performance liquid chromatography system (Wilmington, DE). Before analysis of the sample, the SEC column was first equilibrated with a 20 mM ammonium acetate mobile phase (pH = 7.4) at 0.06 mL/min flow rate for 1 hour. During the analysis, 5 μL of an incubated sample solution was injected into the sample loop. A variable wavelength detector set at 214 nm was used for detection, except in cases where mass spectrometric detection was performed. For the molecular weight (MW) calibration, a solution containing a mixture of the following proteins was used: 1.5 μM bovine serum albumin (MW = 66,000), 3 μM carbonic anhydrase (MW = 29,040), 3 μM myoglobin (MW = 16951), and 3 μM β2m (MW = 11732).

6.2.10 Transmission Electron Microscopy (TEM)

TEM images were obtained on a JEOL JEM-100CX electron microscope operating at 100 keV. Before analysis, a solution containing insoluble material was centrifuged (14,000g) for 8 min and decanted, and the solid material was resuspended in 10 mL of the supernatant. The resulting solution was then applied to carbon-coated copper grids and allowed to air-dry for 5 min. The sample was then stained with 1% phosphotungstic acid, airdried for 12 h, and then viewed.

6.3 Results and Discussion

The Cu(II) and Ni(II) concentrations used for the incubation experiments were calculated from the reported K_d values such that 97% of β2m is bound to the metal of interest. The concentration that was initially used for Zn(II) was also calculated from the reported K_d; however, we observe aggregation/precipitation of β2m when greater than a 5-fold molar excess of Zn(II) is used. Hence, a 4-fold molar excess (400 μM) was chosen for the experiments involving Zn, resulting in 20% loading of the protein by the Zn(II).

6.3.1 Effect on β2m Structure

Assessment of the β2m structure in the presence of the three divalent metal ions was first performed using far- and near-UV CD. All peaks in the far-UV spectrum demonstrate weak negative ellipticity, which is characteristic of a β-sheet fold (Figure 6.1A). The spectrum shows a slightly more negative ellipticity in the presence of Cu(II) and Zn(II) at 220 nm compared to apo-β2m. This may be attributed to the breaking of one or more hydrogen bonds upon binding of these two metals, resulting in small changes in the secondary structure. In the near-UV CD, positive bands were observed at 260 nm and 290 nm for all samples (Figure 6.1B), indicating that the protein is folded and has very similar arrangements of aromatic residues when each of the metals is bound. These bands arise from the packing of aromatic groups into well-defined and fixed conformations. With Cu(II) and Zn(II), there are slight variations from the apo-protein at ~270 and ~ 280 nm, indicating that there might be small changes in the atomic arrangement of the aromatic groups. Overall, the CD data suggest that the global structure of β2m remains mostly unchanged upon interaction with all three divalent metal ions.

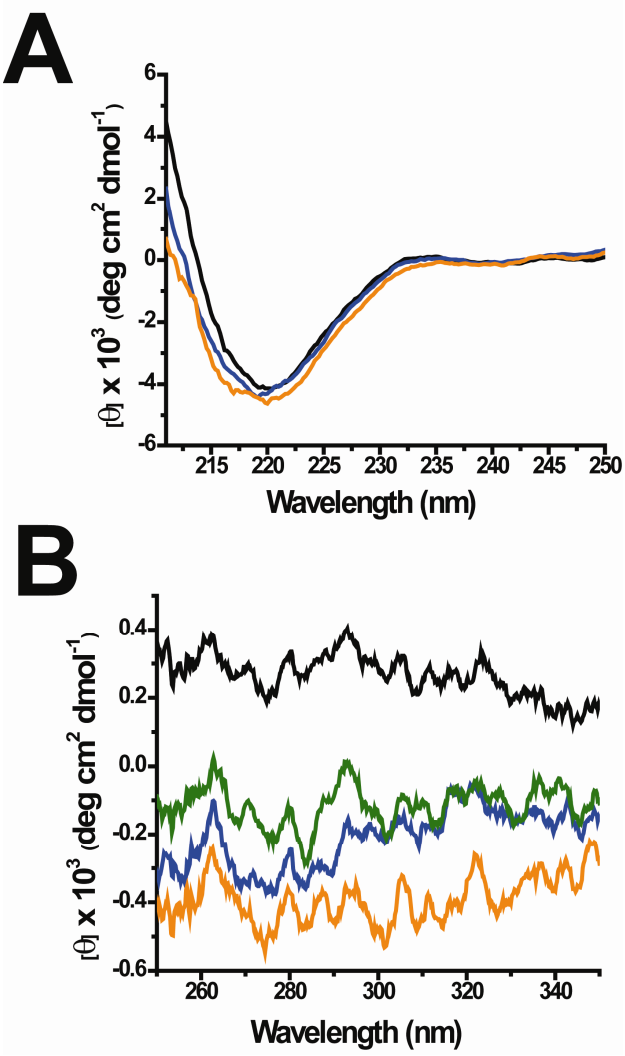


Figure 6.1 CD spectra of β2m in the absence (black) and presence of divalent metal ions: Cu(II) (blue), Ni(II) (green), and Zn(II) (orange). (A) Far-UV CD spectra. (B) Near-UV CD spectra.

6.3.2 Effect on Oligomerization of β2m

The oligomerization of β2m in the presence of the divalent metal ions was monitored using two methods: (1) thioflavin T (ThT) fluorescence and (2) SEC. Upon binding to amyloid-like species, the characteristic fluorescence spectrum of ThT changes

[10]; although, we and others have shown that the spectrum of ThT also changes upon formation of the β2m dimer [8,11,12]. Figure 6.2 shows the ThT fluorescence in the absence and presence of the metal ions. No significant change in ThT fluorescence was found in the control experiment (i.e. no metal ion present) for up to 7 hours. ThT fluorescence in the presence of Ni(II) was similar to that of the control, indicating that oligomerization does not occur with Ni(II), which is consistent with previous studies [13]. In contrast, ThT fluorescence in the presence of Cu(II) and Zn(II) increased immediately after addition of the metal ion. For the Cu-containing solution, the increase in ThT fluorescence enhancement occurs at a rate of 1.38×10^{-1} s^{-1}, and the signal levels off after about 5 hours. In the presence of Zn(II), the ThT fluorescence increases exponentially with a rate of 4.49×10^{-1} s^{-1}, and the signal levels off after 1 hour. This indicates that kinetics of oligomerization is much faster in the presence of Zn(II) compared to Cu(II). Although a change in the fluorescence of ThT is observed within 1 hour for both Cu(II) and Zn(II), no insoluble material is observed from these samples for up to seven hours. Changes in ThT fluorescence by upon binding Cu(II) and Zn(II) indicate that both metals induce the formation of oligomers that are able to bind ThT in a similar manner, whereas Ni either does not induce oligomerization or produces oligomers that are unable to change the fluorescence characteristics of ThT.

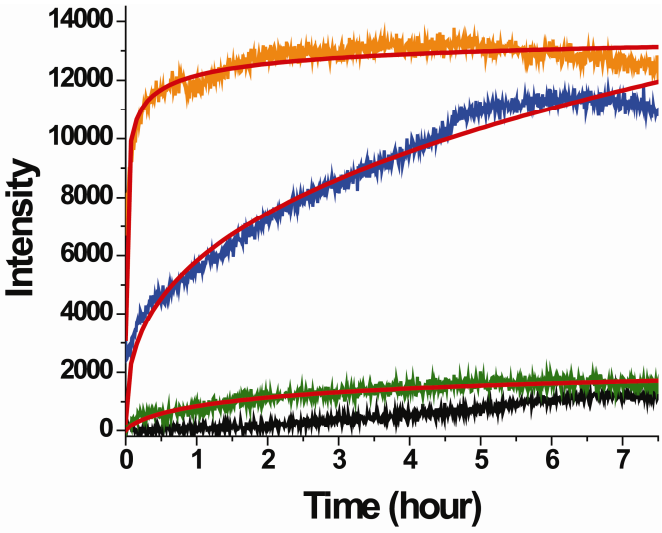


Figure 6.2 Oligomerization of β2m in the absence (black) and presence of divalent metal ions: Cu(II) (blue), Ni(II) (green), and Zn(II) (orange) monitored by thioflavin T (ThT) fluorescence at 483 nm. The ThT fluorescence maximum shifts from 450 nm to 483 nm upon binding to β2m oligomers or amyloid-like species in solution. The trend lines associated with the plots of the metal-containing samples are from non-linear curve fitting of the data.

The oligomerization of β2m was also monitored by SEC in order to follow the kinetics and identities of the oligomeric species that are formed. SEC analyses indicate that no oligomers are formed when β2m is incubated with Ni(II) over the course of two weeks (Figure 6.2). Based on comparison to a calibration curve, the compound that elutes at ~27 min is the monomer. Over the course of two weeks a small drop in the monomer intensity is observed because the protein adheres to the surface of the tubes during incubation over such long periods of time.

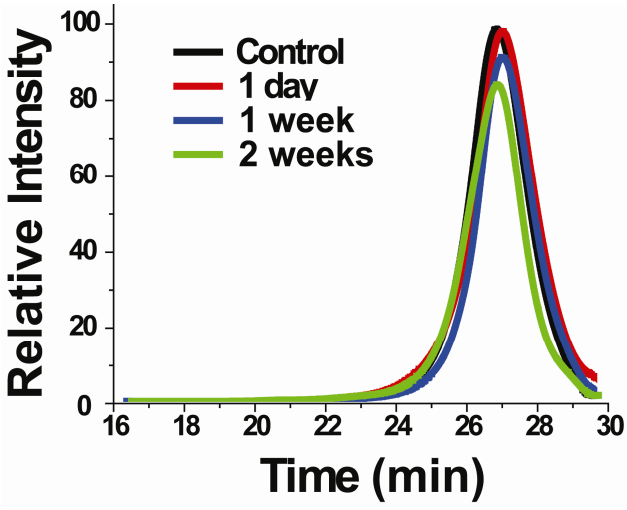


Figure 6.3 β2m oligomerization in the presence of Ni(II) monitored by SEC. The monomer peak is observed at ~27 min.

In the presence of Cu(II), monomers, dimers, tetramers, and hexamers are observed within six days of incubation (Figures 6.4A and 6.4B), which is perfectly consistent with previous studies by our group [12]. Based on a calibration curve, the molecular weights (MW) of the compounds that elute at 20, 21, 25, and 27 min are 11,750 Da, 21,780 Da, 45,240 Da, and 68,310 Da, respectively,. The peak at 27 min with a MW of 11,750 matches very well the expected MW of the monomer (11,731 Da). This assignment was also confirmed by online SEC-MS analysis (data not shown). The expected MW of the dimer (23,462) is close to the MW of the peak at 25 min. Online SEC-MS analysis of this peak confirms that this compound is the β2m dimer (Figure 6.4C). The peak at 21 min has an estimated MW of 45,240 Da which is very close to the expected MW of the tetramer (46,924 Da). The peak that elutes at 20 min is assigned to the hexameric form of β2m, which has an expected MW of 70,386. Unfortunately, the assignments of the tetramer and hexamer peaks could not be confirmed by SEC-MS.

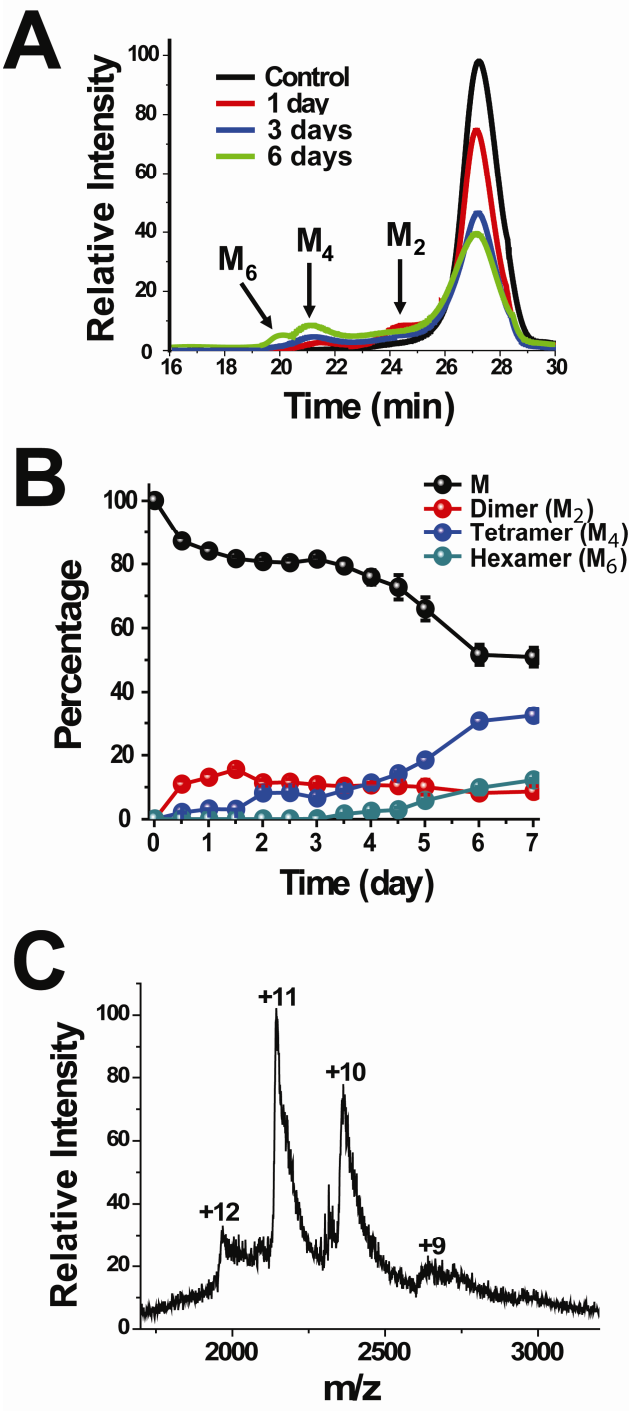


Figure 6.4 SEC and SEC-MS analyses of β2m incubated with Cu(II). (A) Chromatograms for a control (black) and samples incubated with Cu(II) for 1 day (red), 3 days (blue), and 6 days (green). Monomer (M), dimer (M_2), tetramer (M_4), and hexamer (M_6) are observed. (B) Temporal progression of the oligomers measured by SEC. (C) SEC-ESI mass spectrum of the chromatographic peak eluting at ~25 min after 6 days of incubation, confirming that this peak corresponds to the β2m dimer. The charge state of each mass spectral peak is labeled.

Figure 6.5 show that β2m also forms oligomers upon interaction with Zn(II). Unlike with Cu(II), in which there is a gradual decrease in the abundance of the monomer (27 min), the abundance of the monomer in the presence of Zn(II) drops by ~50% after 1 day (Figure 6.5A). This is consistent with the higher rate of oligomerization seen in the ThT fluorescence experiments with Zn(II) (see Figure 6.2). The peak at 23 min has an estimated MW of ~27,000 which is about 3500 more than the expected MW of the dimer but 8000 less than the expected MW of the trimer. SEC-MS analysis confirms that the oligomer eluting at 23 min is the dimer (Figure 6.5C). The oligomer with an elution time of 18 min has a MW of ~79,000 which is about 8600 more than the expected MW of the hexamer and 3000 less than the expected MW of the heptamer. Definitive assignment of this oligomer is difficult because this MW is near the exclusion limit of the SEC column and its identity could not be confirmed by SEC-MS. An interesting observation that is evident in Figure 6.5A is that the retention times of the compounds shift over time. This phenomenon could occur because different oligomers are formed over time or because the oligomers are somewhat unstable and dissociate during the chromatographic process.

Overall, the SEC experiments indicate that the β2m oligomers that are formed in the presence of Cu(II) include dimers, tetramers, and hexamers (Figure 6.4B). The oligomers form via the assembly of dimers since no odd-numbered oligomers were observed. This is consistent with previous SEC and dynamic light scattering experiments [12]. In contrast, β2m immediately forms a dimer upon interaction with Zn(II) and subsequently forms oligomers larger than a hexamer. Thus, in comparison to Cu(II), binding of Zn(II) to β2m causes subtle but different changes in the structure of β2m that result in the formation of slightly different oligomeric forms.

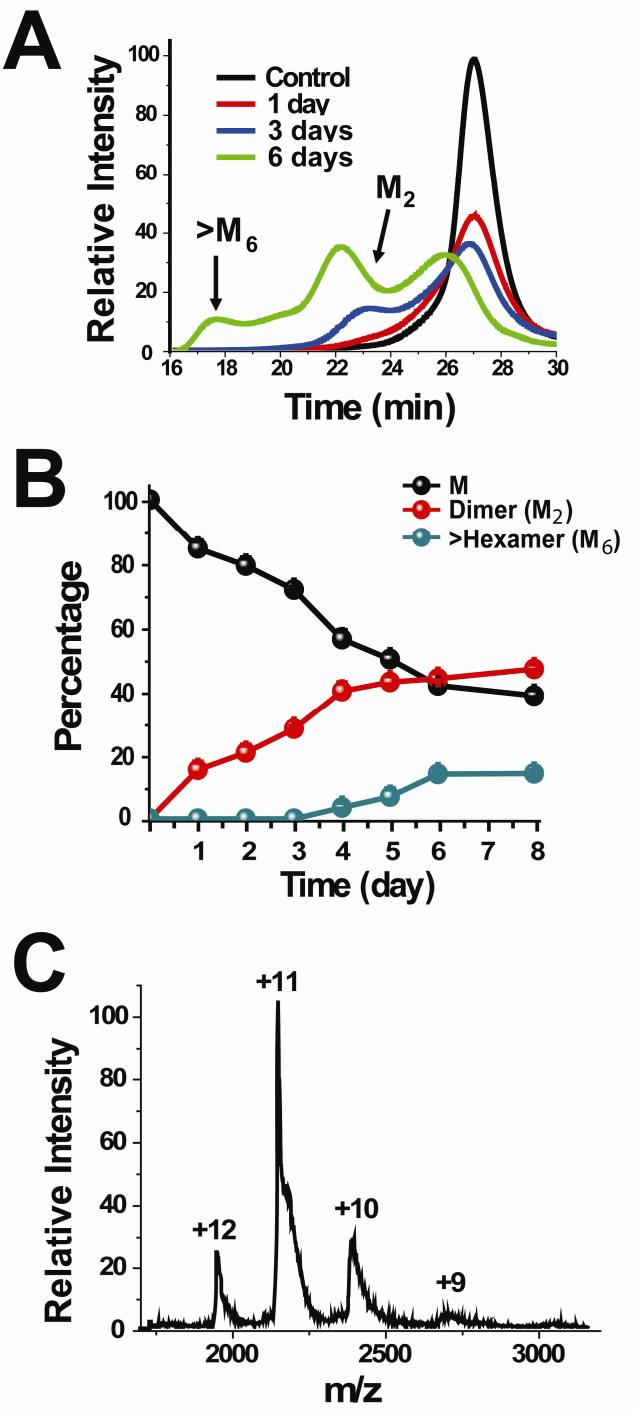


Figure 6.5 SEC and SEC-MS analyses of β2m incubated with Zn(II). (A) Chromatograms for a control (black) and samples incubated for 1 day (red), 3 days ((blue), and 6 days (green). Monomer (M), dimer (M_2), oligomers bigger than a hexamer are observed. (B) Temporal progression of the oligomers measured by SEC. (C) SEC-ESI mass spectrum of the chromatographic peak eluting at ~23 min after 6 days incubation, confirming that this is the β2m dimer. The charge state of each mass spectral peak is labeled.

6.3.3 Effect on Amyloid Formation of β2m

Sodium dodecyl sulfate (SDS) dissolution was used to determine if the aggregates formed in the presence of Cu(II) and Zn(II) are amyloid in nature. One characteristic feature of amyloid-like aggregates as opposed to amorphous aggregates is that they do not readily dissolve in 2% SDS. The aggregates were filtered and added to a 2% solution of SDS at 37°C. The aggregates in the samples incubated with Cu(II) for 4 weeks did not dissolve after 24 h, in the presence of SDS, indicating that they were amyloid-like aggregates. In contrast, the insoluble material formed after 4-week incubation with Zn(II) dissolved in 2% SDS suggesting that they are amorphous aggregates.

We have shown previously that after 1 month of incubation with Cu(II), the vast majority of the aggregates are long and straight, which are hallmarks of amyloid fibrils [12]. On the other hand, TEM images of the aggregates taken after incubation with Zn(II) for 1 month shows that the aggregates formed with Zn(II) do not have well-defined structures (Figure 6.6). Taken as a whole, the SDS dissolution experiments and TEM images suggest that amyloid formation is observed with Cu(II) but not with Zn(II).

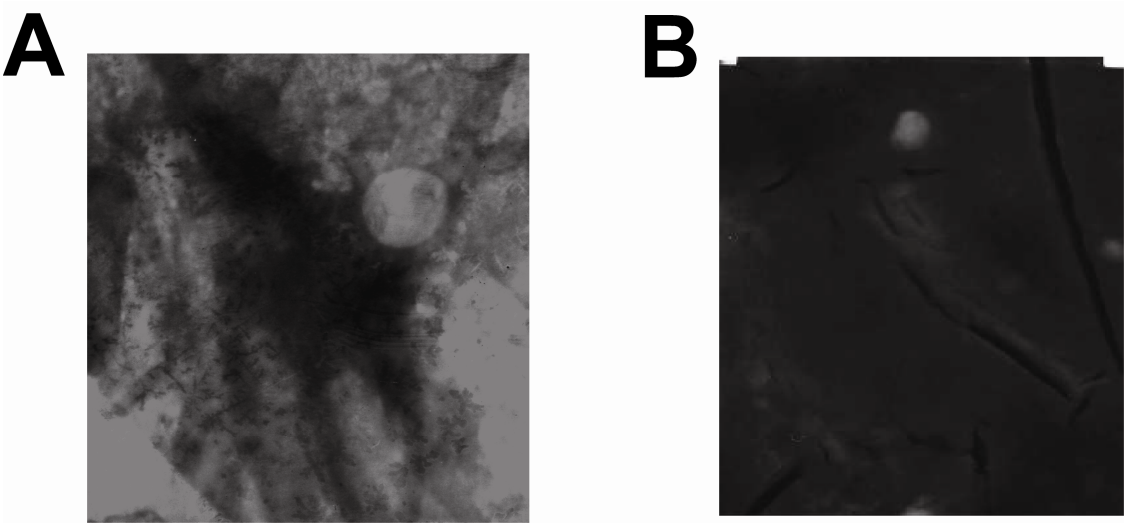


Figure 6.6 (A) Transmission electron micrograph images obtained from β2m after incubation with Zn(II) for 1 month. Dimensions of image is 200 nm x 325 nm. (B) The TEM image for β2m incubated with Cu(II) for 1 month is adapted from reference 12.

6.3.4 Covalent Labeling

The results above reveal that binding to Ni(II) and Zn(II) causes noticeable differences in oligomerization and amyloid formation of β2m when compared to Cu(II) binding. We hypothesize that these differences are due to differences in binding sites and binding-induced conformational changes. The structural changes caused by the binding of each metal ion were investigated by covalent labeling and MS. The reactivity of three covalent labeling reagents with β2m in the presence of each selected metal ion before it formed any oligomers was performed. The three labeling agents that were used (i.e. DEPC, NHSA, and BD) were the same ones described in previous chapters of this dissertation. DEPC can react with solvent exposed His, Tyr, Thr, Lys, and Ser residues [14], but the reactions with this reagent are significantly slowed when these residues, especially His, are bound to metal ions [15,16]. Likewise, NHSA and BD can react with

amino and guanidinium groups, respectively, when these sites are solvent exposed [17], but its reactivity is reduced when these residues are protected, for example, by metal binding. Moreover, increases or decreases in the reactivity of specific amino acids, relative to the apo-protein, can also reflect differences in solvent exposure caused by metal-binding induced conformational changes. All covalent labeling experiments were performed when only monomeric β2m was present. A list of the experimentally-determined modification percentages for each of the modified amino acids reacted in the absence (i.e. apo- β2m) and presence of the divalent metal ions is shown in Table 6.1.

Table 6.1 Modification percentages[a] for each modified amino acid in the absence of metal and the presence of Cu(II), Ni(II), and Zn(II).

Residue	apo-β2m	Cu(II)	Ni(II)	Zn(II)
NHSA				
N-terminus	81 ± 3	60.5 ± 0.7	74 ± 3	71 ± 2
Lys6	20 ± 1	19 ± 1	26.9 ± 0.9	19.1 ± 0.7
Lys19	7.0 ± 0.4	6.9 ± 0.4	5.1 ± 0.4	6.9 ± 0.3
Lys41	13.6 ± 0.7	15 ± 1	13 ± 0.8	13.4 ± 0.9
Lys58	16.1 ± 0.7	16 ± 3	11.8 ± 0.5	14.5 ± 0.6
Lys75	1.6 ± 0.2	1.3 ± 0.1	1.60 ± 0.04	2.3 ± 0.2
Asn83	1.50 ± 0.04	1.4 ± 0.1	1.3 ± 0.2	2.6 ± 0.2
Lys91	27 ± 1	22 ± 1	29 ± 1	24 ± 1
Lys94	3.1 ± 0.2	3.4 ± 0.3	3.8 ± 0.1	4.9 ± 0.2
DEPC				
N-terminus	99 ± 1	68 ± 1	92 ± 1	90 ± 1
Thr4	86.7 ± 0.7	87 ± 2	86 ± 3	86 ± 5
Lys6	7.57 ± 0.05	7.5 ± 0.1	7.93 ± 0.07	7.6 ± 0.1

This table is continued on the next page.

His13	47 ± 2	46 ± 3	46 ± 2	46 ± 3
Lys19	22 ± 1	19 ± 2	18.0 ± 0.7	24 ± 2
Ty26	0.75 ± 0.07	0.79 ± 0.05	1.10 ± 0.05	0.75 ± 0.03
Ser28	0.24 ± 0.01	0.17 ± 0.02	0.35 ± 0.02	0.26 ± 0.01
His31	1.8 ± 0.1	0.7 ± 0.1	1.6 ± 0.1	1.40 ± 0.05
Ser33	1.7 ± 0.1	1.1 ± 0.1	1.11 ± 0.06	1.6 ± 0.1
His51	61 ± 3	61 ± 4	72 ± 2	81 ± 4
Ser57/Lys58	43 ± 2	42 ± 3	19.9 ± 0.8	46 ± 2
Tyr63	7.1 ± 0.5	4.3 ± 0.3	3.6 ± 0.3	5.4 ± 0.3
Tyr67	2.1 ± 0.1	1.75 ± 0.06	2.4 ± 0.1	2.01 ± 0.08
Lys75	0.36 ± 0.04	0.33 ± 0.02	0.35 ± 0.04	0.52 ± 0.04
Ser88	63 ± 3	65 ± 3	60 ± 1	74 ± 1
Lys94	30 ± 2	29 ± 2	39 ± 3	39.7 ± 0.5
BD				
Arg3	0.6 ± 0.1	1.05 ± 0.05	0.48 ± 0.07	0.64 ± 0.04
Arg12	1.3 ± 0.1	1.00 ± 0.05	0.51 ± 0.08	0.9 ± 0.1
Arg45	3.8 ± 0.3	3.6 ± 0.1	3.3 ± 0.1	4.40 ± 0.06
Arg97	2.8 ± 0.4	4.8 ± 0.2	4.9 ± 0.4	1.7 ± 0.4

[a] The percent modification of each labeled amino acid was determined from the ion intensities of both modified and unmodified proteolytic peptide fragments containing the amino acid of interest in a manner identical to that described in Section 3.2.8.

When the covalent labeling reactions were performed 2 minutes after addition of Cu(II), the extent of modification of 10 residues changed when compared to the protein in the absence of any metal. The residues that change in reactivity are localized at the N-terminus region and the ABED sheet (Figure 6.7A). As discussed in Section 3.3.4, the N-terminus and His31 are part of the Cu(II) binding site. Of the 24 modifiable residues in β2m, these two amino acids had the biggest drop in reactivity in the presence of Cu(II).

The other six residues that decrease in reactivity in the presence of Cu(II) are Arg12, Ser28, Ser33, Tyr63, Tyr67, and Lys91. There are only two amino acids that increase in reactivity upon interaction with Cu(II), namely, Arg3 and Arg97. The possible reasons for these structural changes are discussed in Chapter 3.

The reactivity pattern is different when β2m binds Ni(II). Figure 6.7B shows that few residues decrease in reactivity at the N-terminal region of the protein, and several residues increase in reactivity at the ABED β-sheet in the presence of Ni(II). The seven residues that decrease in modification extent after interaction with Ni(II) include the N-terminus, Arg12, Lys19, Ser33, Arg45, Lys58 (and possibly Ser57), and Tyr63. Unlike with Cu(II) where a ~30% decrease in modification extent was observed for the N-terminus, the change in reactivity with Ni(II) was only ~7%. In addition, there is no change in the extent of modification of His31. These results suggest that Ni(II) might bind to a different region of the protein. The extent of modification of Arg12, Lys19, Ser33, Arg45, Lys58, and Tyr63 decrease by 30, 22, 34, 13, 40, and 49%, respectively; however, these residues are not likely to bind Ni(II). Several residues increase in modification extent, namely, Tyr26 (45%), Ser28 (45%), His51 (18%), Lys94 (26%), and Arg97 (97%), indicating that these residues are more accessible after interaction with Ni(II). The conformational changes responsible for these reactivity changes are clearly very different from those observed with Cu(II), which may be the reason why no β2m oligomerization or amyloid formation is observed in the presence of Ni(II).

The amino acids that change in reactivity upon interaction with Zn(II) are shown in Figure 6.7C. Five residues have lower reactivity, namely, the N-terminus, His31, Tyr63, Arg12, and Arg97. The N-terminus and His31 decreased in modification by 11

and 22%, respectively. These modest drops in reactivity may suggest that Zn (II) binds at the N-terminal region in a fashion similar to Cu(II). In contrast to Cu(II), though, Zn(II) binding to the N-terminal region of the protein does not induce any reactivity changes to Ser28, Ser33, and Arg3. The modification extents of Tyr63, Arg12, and Arg97 decrease by 23, 30, and 39%, respectively. In the presence of Cu(II), Tyr63 and Arg12 also decrease in reactivity but the modification extent of Arg97 increased. In addition, we find the levels of modification of Arg45, His51, Lys75, Asn83, Ser88, and Lys94 increase by about 15, 32, 44, 73, 17, and 45%, respectively, in the presence of Zn(II). In contrast, the reactivity of these five residues remain the same in the presence of Cu(II).

Taken together, the covalent labeling data indicates that Cu(II) and Zn(II) may have the same binding site (i.e. N-terminal region), whereas Ni(II) may bind elsewhere. The similarity in binding sites between Cu(II) and Zn(II) might explain why β2m oligomerization is observed for both of these metal ions but not Ni(II). The smaller decrease in modification of the N-terminus and His31 in the presence of Zn(II) may be accounted for by the lower loading of β2m with Zn(II), 20% compared to 97% with Cu(II). Future experiments using EXAFS and metal-catalyzed oxidation (MCO) are necessary to completely delineate the Zn(II) and Ni(II) binding sites to more fully understand the effect of each metal.

The existing covalent labeling data does, however, allow one to speculate about the reasons why Cu(II) induces the formation of amyloid fibrils but Zn(II) does not. Previous studies showed that the structural changes induced by Cu(II) binding enables dimer formation of β2m, which initiates the pathway toward amyloid fibril formation. These structural changes include: (1) drop in reactivity of Ser33 due to *cis-trans*

isomerization of Pro32 [13,14], (2) increased reactivity of Arg3 making salt bridge formation with Glu16 of an adjacent monomer possible [18], and (3) decreased reactivity of Arg12, Ser28, Tyr63, and Tyr67 on the ABED β-sheet which help facilitate formation of a dimer with an antiparallel ABED-ABED interface [19]. These rearrangements are not reflected in the covalent labeling data in the presence of Zn(II). For instance, Ser33 has the same modification extent in the absence and presence of Zn(II). This indicates that this residues does not undergo any rearrangement and that the *cis* to *trans* backbone isomerization of Pro32 does not likely occur upon interaction with Zn(II). Likewise, the reactivity of Arg3 did not change suggesting that it has the same conformation in the apo and Zn(II)-bound β2m. Salt bridge formation with Glu16 of an adjacent β2m monomer is probably not favored by this conformation. In addition, only Arg12 and Tyr63 on the ABED β-sheet decreased in reactivity and several residues on the CFG β-sheet increased in reactivity upon interaction with Zn(II). These rearrangements may likely facilitate formation of a dimer with an interface distinct from that present in the dimer formed with Cu(II). This alternative dimer does not form tetramers; it may have rearrangements that give rise to the formation of hexameric interfaces instead. Future covalent labeling experiments on this alternative dimer is required to determine its interface and any alterations that facilitate hexamer formation. Moreover, Cu(II) and Ni(II), but not Zn(II), are known to be able to bind to backbone amides [18,20-26]. We speculate that Cu(II) binding to the backbone amide between Ile1 and Gln2 is important for causing the structural changes necessary for β2m to form amyloid-competent oligomers. Hence, these differences in metal-binding induced structural changes possibly account for the fact that amyloid formation of β2m is only observed with Cu(II).

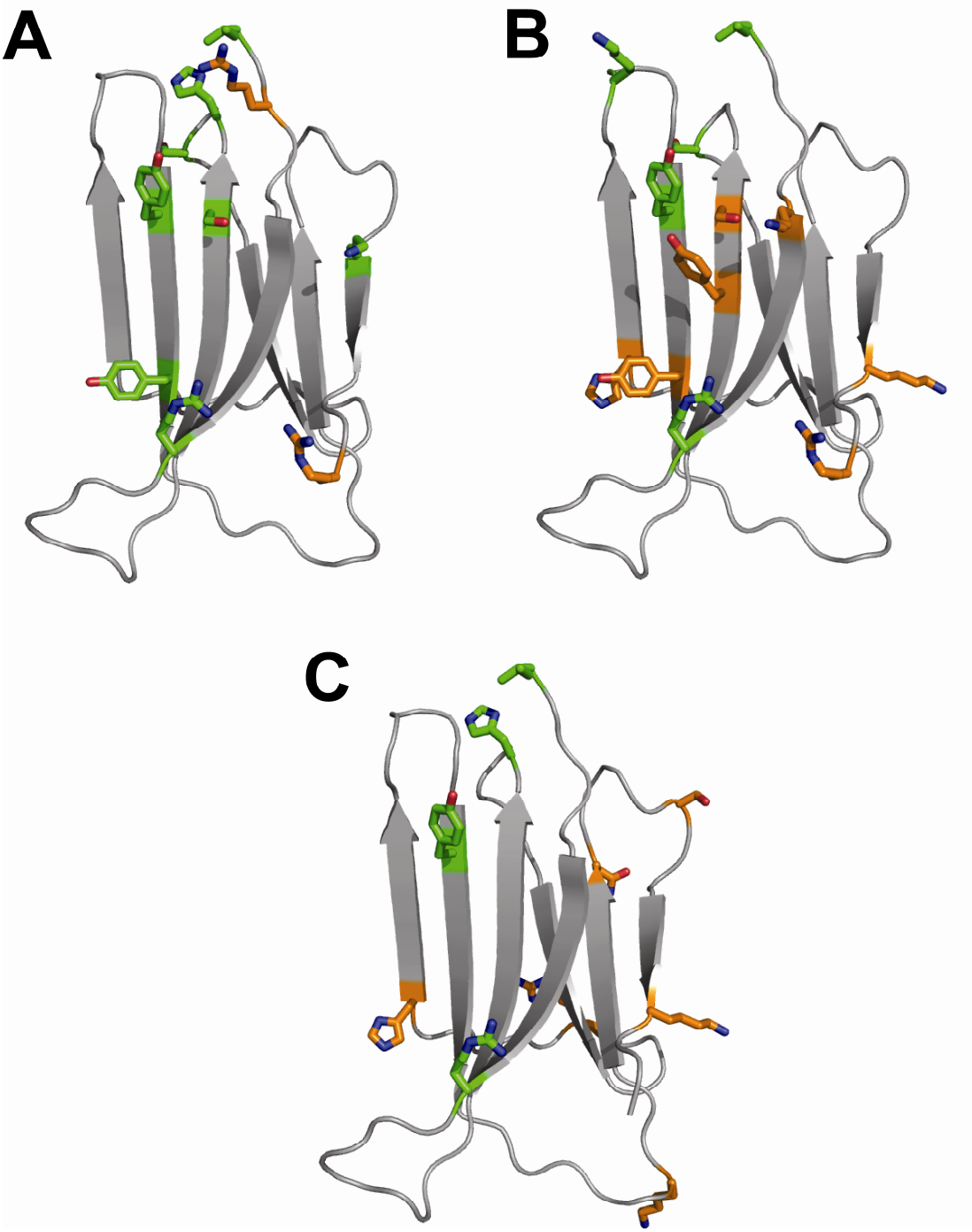


Figure 6.7 Ribbon representation of monomeric β2m (PDB ID: 2D4F). Amino acids that decrease (green) and increase (orange) in reactivity upon interaction with divalent metal ions are shown as sticks. (A) Cu(II). (B). Ni(II). (C). Zn(II).

6.4 Conclusions

Cu(II) but not Zn(II) or Ni(II) can induce the amyloid formation of β2m. The unique effect of Cu(II) appears to be caused by the site at which the metal binds the protein and the conformational changes it induces. Ni(II) does not cause oligomerization or amyloid formation because it binds at a different site on the protein than Cu(II) does. Zn(II) causes oligomerization and aggregation but does not induce amyloid fibril formation. Zn(II) binds in the N-terminal region and is bound by the N-terminus and His31. However, we speculate that Zn(II) is not bound to an amide backbone like Cu(II). This difference possibly induces structural changes in β2m that are different from those observed in the presence of Cu(II). The dimer formed with Zn(II) may therefore be different from the antiparallel ABED-ABED arrangement observed with Cu(II). We propose that the rearrangements in this alternative dimer make formation of hexamers more favorable. Lastly, covalent labeling along with MS detection appears to be a useful method for obtaining amino acid-level information about metal binding-induced conformational changes in proteins. This ability could make this method useful for other amyloid forming systems in which binding of metal ions have been implicated.

6.5 References

1. Jobling, M.F., Huang, X., Stewart, L.R., Barnham, K.J., Curtain, C., Volitakis, I., Perugini, M., White, A.R., Cherny, R.A., Masters, C.L., Barrow, C.J., Collins, S.J., Bush, A.I., and Cappai, R. (2001) Copper and zinc binding modulates the aggregation and neurotoxic properties of the prion peptide PrP106-126. *Biochemistry 40*, 8073-8084.

2. Bush, A.I., and Tanzi, R.E. (2002) The galvanization of beta-amyloid in Alzheimer's disease. *Proc. Natl. Acad. Sci. USA 99*, 7317-7319.

3. Zatta, P., Drago, D., Bolognin, S., and Sensi, S.L. (2009) Alzheimer's disease, metal ions and metal homeostatic therapy. *Trends Pharmacol. Sci. 30*, 346-355.

4. Cole, N.B., Murphy, D.D., Lebowitz, J., Di Noto, L., Levine, R.L., and Nussbaum, R.L. (2005) Metal-catalyzed oxidation of alpha-synuclein: helping to define the relationship between oligomers, protofibrils, and filaments. *J. Biol. Chem.* *280*, 9678–9690.

5. Uversky, V.N., Li, J., and Fink, A.L. Metal-triggered structural transformations, aggregation, and fibrillation of human alpha-synuclein. (2001) A possible molecular NK between Parkinson's disease and heavy metal exposure. *J. Biol. Chem. 276*, 44284-44296.

6. Davis, D.P., Gallo, G., Vogen, S., Dul, J.L., Sciarretta, K.L., Kumar, A., Raffen, R., Stevens, F.J., and Argon, Y. (2001) Both the environment and somatic mutations govern the aggregation pathway of pathogenic immunoglobulin light chain. *J. Mol. Biol. 313*, 1021-1034.

7. Eakin, C.M., Knight, J.D., Morgan, C.J., Gelfand, M.A., and Miranker, A.D. (2002) Formation of a copper specific binding site in non-native states of β-2-microglobulin. *Biochemistry 41*, 10646-10656.

8. Eakin, C.M., Attenello, F.J., Morgan, C.J., and Miranker, A.D. (2004) Oligomeric assembly of native-like precursors precedes amyloid formation by β-2-microglobulin. *Biochemistry 43*, 7808-7815.

9. Morgan, C.J., Gelfand, M., Atreya, C., and Miranker, A.D. (2001) Kidney dialysis-associated amyloidosis: A molecular role for copper in fiber formation. *J. Mol. Biol. 309*, 339-345.

10. Naiki, H., Hasegawa, K., Yamaguchi, I., Nakamura, H., Gejyo, F., and Nakakuki, K. (1998) Apolipoprotein E and antioxidants have different mechanisms of inhibiting Alzheimer's beta-amyloid fibril formation in vitro. *Biochemistry 37*, 17882-17889.

11. Calabrese, M. F. and Miranker, A. D. (2007) Formation of a stable oligomer of β-2 microglobulin requires only transient encounter with Cu(II). *J. Mol. Biol. 367*, 1-7.

12. Antwi, K., Mahar, M., Srikanth, R., Olbris, M. R., Tyson, J. F., and Vachet, R. W. (2008) Cu(II) organizes beta-2-microglobulin oligomers but is released upon amyloid formation. *Protein Sci. 17*, 748-759.

13. Eakin, C.M., Berman, A.J., and Miranker, A.D. (2006) A native to amyloidogenic transition regulated by a backbone trigger. *Nature Struct. Mol. Biol. 13*, 202-208.

14. Mendoza, V. L., and Vachet, R. W. (2008) Protein surface mapping using diethylpyrocarbonate with mass spectrometric detection. *Anal. Chem. 80*, 2895-2904.

15. Qin K, Yang Y, Mastrangelo P, Westaway D. (2002) Mapping Cu(II) binding sites in prion protein by diethyl pyrocarbonate modification and matrix-assisted laser desorption ionization-Time of Flight (MALDI-TOF) mass spectrometric footprinting. *J. Biol. Chem. 277*, 1981-1990.

16. Narindrasorasak S, Kulkarni P, Deschamps P, She Y, Sarkar B. (2007) Characterization and copper binding properties of human COMMD1 (MURR1). *Biochemistry 46*, 3116-3128.

17. Mendoza, V. L., and Vachet, R. W. (2009) Probing protein structure by amino acid-specific covalent labeling and mass spectrometry. *Mass Spec. Rev. 28*, 785-815.

18. Srikanth, R., Mendoza, V. L., Bridgewater, J. D., Zhang, G., and Vachet, R. W. (2009) Copper binding to β-2-microglobulin and its pre-amyloid oligomers. *Biochemistry 48*, 9871-9881.

19. Mendoza, V.L., Antwi, K., Baron-Rodriguez, M.A., Blanco, C., and Vachet, R.W. (2009) Structure of the preamyloid dimer of beta-2-microglobulin from covalent labeling and mass spectrometry. *Biochemistry 49*, 1522-1532.

20. Perkins, C.M., Roes, N.J., Weinstein, B., Stenkamp, R.E., Jensen, L.H., and Pickart, L. *(*1984*)* The structure of a copper complex of the growth-factor glycyl-L-histidyl-L-lysine at 1.1 A resolution. *Inorg. Chim. Acta* 82, 93–99.

21. Camerman, N., Camerman, A., and Sarkar, B. *(*1976*)* Molecular design to mimic the copper(II) transport site of human albumin. The crystal and molecular structure of copper(II) – glycylglycyl-L-histidine-*N*-methyl amide monoaquo complex. *Can. J. Chem. 54*, 1309–1316.

22. Aronoff-Spencer, E., Burns, C.S., Avdievich, N.I., Gerfen, G.F., Peisach, J., Antholine, W.E., Bal, H.L., Cohen, F.E., Prusiner, S.B., and Millhauser, G.L. (2000) Identification of the Cu^{2+} binding sites in the N-terminal domain of the prion protein by EPR and CD spectroscopy. *Biochemistry 39*, 13760–13771.

23. Xia, W., Li, H.Y., Sze, K.H., and Sun, H.Z. (2009) Structure of a nickel chaperone, HypA, from *Helicobacter pylori* reveals two distinct metal binding sites. *J. Am. Chem. Soc. 131*, 10031-10040.

24. Van Horn, J.D., Bulaj, G., Goldenberg, D.P., and Burrows, C. J. J. (2003) The Cys-Xaa-His metal-binding motif: {N} versus {S} coordination and nickel-mediated formation of cysteinyl sulfinic acid. *Biol. Inorg. Chem. 8*, 601– 610

25. Wuerges, J., Lee, J.W., Yim, Y.I., Yim, H.S., Kang, S.O., and Carugo, K.D. (2004) Crystal structure of nickel-containing superoxide dismutase reveals another type of active site. *Proc. Natl. Acad. Sci. USA 101*, 8569– 8574.

CHAPTER 7

SUMMARY AND FUTURE DIRECTIONS

7.1 Summary

This dissertation has described the development of a protein surface mapping method combined with mass spectrometry to elucidate the structural features of the oligomers that precede Cu(II)-induced β-2-microglobulin (β2m) amyloid fibril formation. The method involves covalent modification of amino acid side chains that are exposed to solvent and the identification of the modified residues by MS. This method was found to be very powerful for monitoring conformational changes in β2m as well as identifying the protein-protein interfaces of its pre-amyloid dimer and tetramer.

The improved covalent modification and MS-based method addresses the need for a reliable approach that both ensures protein structural integrity during labeling experiments and provides readily detectable modifications. We demonstrated that measuring the kinetics of the covalent labeling reactions ensures that the covalent probe does not disrupt a protein's structure during the labeling reaction. The dose-response plots provide the added benefit of quantitative information about residue reactivity, which could possibly provide more precise structural information. The modification rate coefficients seem to reflect both the reactive atom's solvent accessibility and amino acid microenvironment, but stronger connections between these parameters requires further work. In addition, we show that DEPC may have promise as a general covalent label because of its ability to probe up to 25% of the residues in the average protein.

Covalent labeling in conjunction with mass spectrometry was used to determine the Cu(II) binding site in β2m and conformational changes induced by metal binding. The labeling data confirms the involvement of the N-terminal amine and His31 in the binding site and further rules out the interaction of His51 with Cu(II). The reactivity patterns in the absence and presence of Cu(II) indicate that the Cu(II)-binding induced conformational changes are predominantly localized at the N-terminal region and the ABED β-sheet. Cu(II) binding to monomeric β2m repositions Arg3 and forces it to adopt a new position that allows it to readily form a dimer-stabilizing intermolecular salt bridge. The rearrangements of Arg12, Ser28, Tyr63, and Tyr67 on the ABED β-sheet possibly creates a hydrophobic face that can promote dimerization via interaction with another hydrophobic ABED β-sheet of an adjacent monomer.

Using three complementary covalent labels with MS detection, we were able to determine models for the dimer and tetramer that precede fibril formation. The dimer was found to have an interface that involves the antiparallel arrangement of ABED sheets from two monomers. Our data, therefore, rules out two competing dimer models - the dimer interface present in the P32A mutant and the interface formed by the C and D chains of the H13F hexamer. Moreover, our covalent labeling data allows us to develop, for the first time, a model for the tetramer in which the interface is mediated by interactions between D strands of one dimer unit and the G strands of another dimer unit. Our results rule out the possibility of a tetramer mediated by interactions between D strands on separate dimer units as suggested previously.

Finally, this surface mapping method was used to obtain insight into the unique effect of Cu(II) on β2m. Cu(II) but not Zn(II) or Ni(II) can induce amyloid formation of β2m. The unique effect of Cu(II) appears to be caused by the site at which the metal binds the protein and the conformational changes it induces. Cu(II) and Zn(II) seem to bind in similar regions of the protein (i.e. N-terminus and His31), whereas Ni(II) may bind elsewhere. The similarity in binding by Cu(II) and Zn(II) might explain why β2m oligomerization is observed for both of these metal ions but not Ni(II). The identity of the other ligands that bind Zn(II) may be different from those with Cu(II) and certainly the binding geometry of the two metals will be different given Zn(II) preference for a tetrahedral binding geometry. Binding to these ligands probably induces structural changes in β2m that are different from those observed in the presence of Cu(II), which is why amyloid formation is not observed with Zn(II).

The covalent labeling of wild-type β2m under amyloid fibril-forming conditions provides structural information relevant for understanding the Cu(II)-induced oligomeric assembly. The knowledge gained in these studies about the subtle structural changes in human β2m required for oligomerization and the interfaces in the dimer and tetramer will serve as reference points in characterizing the larger oligomers (i.e. Cu(II)-free tetramer and hexamer). Further structural changes are expected for the structures reported here to adopt irreversible oligomeric conformations. One possibility is the release of Cu(II) in tetramer I induces subtle structural changes in other regions of the protein but the interfaces reported here are mostly retained. This structural information can also serve as important foundational information in delineating the mechanism of conversion of the oligomers to the insoluble, protease-resistant cross-β structure present in the fibrils.

In addition, this work may contribute significantly not only in detecting relevant targets for diagnosis but also in designing functional therapeutics to treat DRA. Based on the structural information we obtained from the protein surface mapping methods, there are two possible points in the oligomerization process that may be particularly attractive targets for therapeutics. The first point is the conformational changes of the monomer after interaction with Cu(II) that initiate dimer formation. Therapeutic agents can stabilize the monomeric form of the protein, resulting in complete blockage of dimerization and oligomerization. Interestingly, we find that the A and B β-strands of β2m which undergo conformational changes after Cu(II) binding and form part of the dimer interface also comprise the regions of β2m that form non-covalent interactions with the α1 domain of the polymorphic α-chain in the MHC-I complex [1]. The sequence of the α1 domain that makes contact with β2m can be used as reference in designing inhibitors that can bind to monomeric β2m, and thus prevent dimerization. The second point would be the prevention of tetramer formation by stabilization of the β2m dimer. The D β-strand which is an integral part of the proposed tetramer interface also spans the region that interacts with the α3 domain of the MHC-I complex [1]. The amino acid sequence of the α3 domain involved in the noncovalent interaction should permit the design of molecules that prevent β2m oligomerization and fibril formation.

This study also highlights the use of selective covalent modification with mass spectrometric detection in the characterization of metal binding and binding-induced conformational changes as well as protein-protein interactions. This approach has been useful in probing protein structure in situations where more than one protein species is present in solution. This ability makes this method useful for other amyloid forming

systems in which discrete oligomers precede fibril formation. This approach could take on greater importance since in several amyloid diseases the pre-amyloid oligomers may represent the species responsible for cellular toxicity rather than the amyloid fibrils themselves [2,3]. Characterization of the structure of these intermediate oligomers may be crucial to designing novel therapeutic strategies to prevent formation of the oligomer, inhibiting their toxic activity, and in deciphering the mechanism of toxicity.

Furthermore, the common characteristics exhibited by the fibrils from different precursor proteins suggest a general mechanism for amyloid formation [4]. This offers the prospect that the Cu(II)-β2m can be used as a model system to systematically assess the factors that predispose a native protein to form amyloid fibrils. Determining the mechanism of Cu(II)-mediated oligomeric assembly is therefore important to develop a more detailed understanding of the origin and progression of the increasing number of human disorders associated with amyloid formation.

7.2 Future Directions

The following sections discuss future improvements to the covalent labeling method, structural characterization of the β2m oligomers in the presence of Cu(II), and structural characterization of the oligomers formed with Zn(II).

7.2.1 Improving the Protein Surface Mapping Method

7.2.1.1 Multi-Amino Acid Labeling

Simultaneous labeling of numerous amino acids, rather than in parallel, will allow investigation of several sites of the protein structure and thus provide more information per experiment as well as possibly simplifying the modification protocol. Multi-amino acid labeling will entail addition of at least 3 residue-specific labels followed by proteolytic digestion and LC-MS analysis. To date, there has been no protein surface mapping studies using multi-amino acid labeling by selective reagents. Hence, several key issues have to be addressed. Will the reagents react with one another? Can we find reaction conditions that will work for multiple reagents? Will the presence of multiple reagents affect the labeling efficiency of each label? Will using multiple reagents affect the maximum concentration that can be used for a given reagent before protein structure is compromised? How will the amount of modification be accurately determined if a given fragment has two or more different modifications?

7.2.1.2 Minimizing Reversibility of Covalent Modifications

Some covalent modifications are reversible. For instance, mono-modification of histidine residues with DEPC is reversible under both acidic and alkaline conditions and in the presence of nucleophiles such as hydroxylamine and tris(hydroxymethyl)aminomethane (TRIS) [5]. The half-life of histidine residues modified by N-carbethoxyimidazole is about 55 hours at pH 7, 2 hours at pH 2, and 18 min a pH 10 [6]. The adduct formed by arginine with butanedione is unstable at pH < 9 [7]. The reversibility and instability makes detection of modified residues difficult.

One way to minimize these reversible reactions is to decrease the proteolysis time. Natively-folded proteins are known to require many hours for adequate proteolysis by enzymes under conventional conditions. Loss of modifications likely occurs during these long digestion times. Studies have shown that microwave [8,9], ultrasonic [10,11], and immobilized-trypsin [12,13] digestions can significantly accelerate the digestion process to minutes versus hours by traditional methods. With these methods the MS analysis can be performed within a few hours after the modification reactions, and detection of the modified peptides will be easier if they are present in greater amount. However, factors such as digestion completeness and effects of these digestion procedures on modification stability have to be investigated.

Another way to minimize loss of modifications is by using top-down sequencing, which entails performing MS/MS on the whole protein. The key advantage of the top-down approach is that the long enzymatic digestion step can be avoided, which allows the analysis to be done more quickly and possibly avoids the loss of covalent labels that might not be stable during the relatively long digestion step. Top-down sequencing has been used effectively for mapping labeled sites so that 3D protein structural information can be obtained [14-16].

7.2.2 Structural Characterization of Cu(II)-Induced Oligomer Assembly of β2m

In addition to the covalent labels used in this dissertation, other approaches can be used to gain deeper insight into the Cu(II)-induced conformational changes and the structural properties of the pre-amyloid oligomers of β2m. One example is amide H/D exchange, which can be complementary to covalent labeling because it provides

backbone information whereas covalent labels provide information about amino acid side chains. It can potentially provide information that will support the proposed models for the β2m dimer and tetramer. Something similar was done in the study of double-stranded telomeric DNA interactions with the DNA-binding domain of human telomeric repeat binding factor 2 (hTRF2) using selective acetylation of lysine residues in combination with PLIMSTEX (protein-ligand interactions by mass spectrometry, titration, and H/D exchange) [17].

Another covalent label that can potentially provide important structural information is 1-ethyl-3-(3-dimethylaminopropyl)carbodiimide hydrochloride (EDC) combined with nucleophiles like glycine ethyl ester (GEE) and glycineamide (GA). These combined reagents modify acidic amino acid side chains. Reactions with carbodiimides are typically carried out at room temperature in the pH range of 4.5 to 6 [18]. The relatively low pH conditions used for the reactions with EDC make it difficult to use this reagent to probe protein structure under native conditions. However, a recent study showed that a combination of EDC/GEE provides readily detectable modifications at neutral pH [19]. Modification of Cu(II)-bound β2m and the oligomeric intermediates can provide support to our conclusion that Cu(II) binds to the N-terminal region and to our proposed structural models for the pre-amyloid dimer and tetramer. Specifically, previous work demonstrated that Asp59 is one of the Cu(II) binding sites in β2m [20]. Thus, we expect its reactivity with EDC/GEE to decrease in the presence of Cu(II). In addition, we recently postulated that Asp59 and Glu16 play crucial roles in forming electrostatic interactions that stabilize the β2m dimer [21]. We predict that that modification extent for these two acidic residues will decrease upon dimer formation.

7.2.3 Structural Characterization of Zn(II)-Induced Oligomer Assembly of β2m

7.2.3.1 Zn(II) Binding Site

Aside from covalent labeling, the Cu(II) binding site was determined using another MS-based method, metal-catalyzed oxidation (MCO). However, this MCO/MS approach will not be directly applicable to the determination of the Zn(II) binding site because Zn(II) does not undergo single-electron redox processes. One alternative that can identify the ligands that bind Zn(II) is extended X-ray absorption fine structure (EXAFS). EXAFS provides information about the number and type of atoms and their bond distances from a metal atom in a metalloprotein. The X-ray spectra close to the absorption edge of the core electrons show small oscillations, which are caused by scattering of the photoelectron by the surrounding atoms. Therefore, these oscillations contain information about the distance between the metal and the neighboring atoms, the number of neighboring atoms at a certain distance, and the atomic number of these atoms [22]. It is important to determine the Zn(II) binding site to understand more deeply why Zn(II) does not induce amyloid formation even though it also binds in the N-terminal region of β2m.

7.2.3.2 Oligomeric Intermediates

In addition to the Zn(II) binding site, the structural characterization of the oligomers formed by β2m after incubation with Zn(II) is important to fully understand why amyloid formation is only observed in the presence of Cu(II). Covalent labeling with MS detection used to determine the interface of the dimer and tetramer formed with Cu(II) could also be applied to Zn(II)-β2m oligomers. The modification patterns in

different conditions (i.e. Zn(II)-bound monomer, Zn(II)-bound monomer plus dimer, etc.) will help identify the amino acids involved in the oligomer formation. We speculate that with Zn(II), β2m may form a hexamer with interfaces similar to that in the crystal structure of the H13F mutant. This β2m mutant forms hexamer readily but no amyloid fibrils are formed. Therefore, the H13F and Zn(II)-bound β2m both contain structural features that prevent further aggregation. Lastly, the specific role of Zn(II) in the formation and stability of the oligomeric intermediates is important and thus have to be investigated. In the Cu(II)-β2m system, Cu(II) is only necessary for nucleation and its release is required to form amyloid-competent oligomers [23]. One possible scenario is that Zn(II) may be necessary for the stability of all oligomers formed, which may preclude β2m from undergoing the structural changes required to adopt irreversible oligomeric conformations.

7.3 References

1. Bjorkman, P.J., Saper, M.A., Samraoui, B., Bennett, W.S., Strominger, J.L., and Wiley, D.C. (1987) Structure of the human class-I histocompatibility antigen, HLA-A2. *Nature 329*, 506-512.

2. Bucciantini, M., Giannoni, E., Fabrizio, C., Baroni, F., Formigli, L., Zurdo, J., Taddei, N., Ramponi, G., Dobson, C.M., and Stefani, M. (2002) Inherent toxicity of aggregates implies a common mechanism for protein misfolding diseases. *Nature 416*, 507-511.

3. Mendes Sousa, M., Cardoso, I., Fernandes, R., Guimaraes, A., and Saraiva, M.J. (2001) Deposition of transthyretin in early stages of familial amyloidotic polyneuropathy. Evidence for toxicity of nonfibrillar aggregates. *Am. J. Pathol. 159*, 1993-2000.

4. Rochet, J.C., and Lansbury, P.T. Jr. (2000) Amyloid fibrillogenesis: themes and variations. *Curr. Opin. Struct. Biol. 10*, 60-68.

5. Miles, E.W. (1977) Modification of histidyl residues in proteins by diethylpyrocarbonate. *Method Enzymol. 47*, 431-442.

6. Melchoir, W.B. Jr., and Fahrney, D. (1970) Ethoxyformylation of proteins. Reaction of ethoxyformic anhydride with α-chymotrypsin, pepsin, and pancreatic ribonuclease at pH 4. *Biochemistry 9*, 251-258.

7. Riordan, J.F. (1973) Functional arginyl residues in carboxypeptidase-A – Modification with butanedione. *Biochemistry 12*, 3915-3923.

8. Pramanik, B.M., Mirza, U.A., Ing, Y.-H., Liu, Y.-H., Bartner, P.L., Weber, P.C., and Bose, A.K. (2002) Microwave-enhanced enzyme reaction for protein mapping by mass spectrometry: A new approach to protein digestion in minutes. *Protein Sci. 11*, 2676-2687.

9. Sun, W., Gao, S., Wang, L., Chen, Y., Wu, S., Wang, X., Zheng, D., and Gao, Y. (2005) Microwave-assisted protein preparation and enzymatic digestion in proteomics. *Mol. Cell. Proteomics 5*, 769-776.

10. Rial-Otero, R., Carreira, R.J., Cordeiro, F.M., Moro, A.J., Santos, H.M., Vale, G., Moura, I., and Capelo, J.L. (2007) Ultrasonic assisted protein enzymatic digestion for fast protein identification by matrix-assisted laser desorption/ionization time-of-flight mass spectrometry Sonoreactor versus ultrasonic probe. *J. Chromatogr. A* 1166, 101-107.

11. Santos, H.M., Mota, C., Lodeiro, C., Moura, I., Isaac, I., and Capelo, J.L. (2008) An improved clean sonoreactor-based method for protein identification by mass spectrometry-based techniques. *Talanta 77*, 870-875.

12. Jeng, J., Lin, M.-F., Cheng, F.-Y., Yeh, C.-S., and Shiea, J. (2007) Using high-concentration trypsin-immobilized magnetic nanoparticles for rapid in situ protein digestion at elevated temperature. *Rapid Comm. Mass Spectrom. 21*, 3060-3068.

13. Massolini, G., and Calleri, E. (2005) An improved clean sonoreactor-based method for protein identification by mass spectrometry-based techniques. *J. Sep. Sci. 28*, 7-21.

14. McLafferty, F.W., Breuker, K., Jin, M., Han, X.M., Infusini, G., Jiang, H., Kong, X.L., and Begley, T.P. (2007) Top-down MS, a powerful complement to the high capabilities of proteolysis proteomics. *FEBS J. 274*, 6256-6268.

15. Meng, F.Y., Forbes, A.J., Miller, L.M., and Kelleher, N.L. (2005) Detection and localization of protein modifications by high resolution tandem mass spectrometry. *Mass Spectrom. Rev. 24*, 126-134.

16. Novak, P., Kruppa, G.H., Young, M.M., and Schoeniger, J. (2004) A top-down method for the determination of residue-specific solvent accessibility in proteins. *J. Mass Spectrom. 39*, 322-328.

17. Sperry, J.B., Shi, X., Rempel, D.L., Nishimura, Y., Akashi, S., and Gross, M.L. (2008) A mass spectrometric approach to the study of DNA-binding proteins: interaction of human TRF2 with telomeric DNA. *Biochemistry 47*, 1797-1807.

18. Mendoza, V.L., and Vachet, R.W. (2009) Probing protein structure by amino acid-specific covalent labeling and mass spectrometry. *Mass Spectrom. Rev. 28*, 785-815.

19. Wen, J.Z., Zhang, H., Gross, M.L., and Blankenship, R.E. (2009) Membrane orientation of the FMO antenna protein from *Chlorobaculum tepidum* as determined by mass spectrometry-based footprinting. *Proc. Natl. Acad. Sci. USA 106*, 6134-6139.

20. Srikanth, R., Mendoza, V. L., Bridgewater, J. D., Zhang, G., and Vachet, R. W. (2009) Copper binding to β-2-microglobulin and its pre-amyloid oligomers. *Biochemistry 48*, 9871-9881.

21. Mendoza, V.L., Antwi, K., Baron-Rodriguez, M.A., Blanco, C., and Vachet, R.W. (2009) Structure of the pre-amyloid dimer of beta-2-microglobulin from covalent labeling and mass spectrometry. *Biochemistry 49*, 1522-1532.

22. Kodre, A., Arcon, I., and Gomilsek, J.P. (2004) X-ray absorption spectroscopy and related techniques. *Acta Chim. Slovenica 51*, 1-10.

23. Antwi, K., Mahar, M., Srikanth, R., Olbris, M. R., Tyson, J. F., and Vachet, R. W. (2008) Cu(II) organizes beta-2-microglobulin oligomers but is released upon amyloid formation. *Protein Sci. 17*, 748-759.

BIBLIOGRAPHY

Akashi, S., Shirouzu, M., Terada, T., Ito, Y., Yokohama, S., and Takio, K. (1997) Characterization of the structural difference between active and inactive forms of the Ras protein by chemical modification followed by mass spectrometric peptide mapping. *Anal. Biochem. 248*, 15-25.

Antwi, K., Mahar, M., Tyson, J.F., and Vachet, R.W. (2008) Cu(II) organizes β-2-microglobulin oligomers but is released before amyloid formation. *Protein Sci. 17*, 748-759.

Aronoff-Spencer, E., Burns, C.S., Avdievich, N.I., Gerfen, G.F., Peisach, J., Antholine, W.E., Bal, H.L., Cohen, F.E., Prusiner, S.B., and Millhauser, G.L. (2000) Identification of the Cu^{2+} binding sites in the N-terminal domain of the prion protein by EPR and CD spectroscopy. *Biochemistry 39*, 13760–13771.

Berendsen, H., van der Spoel, D., and van Drunen, R. (1995) Gromacs: A message-passing parallel molecular dynamics implementation. *Comp. Phys. Comm. 91*, 43-56.

Bjorkman, P.J., Saper, M.A., Samraoui, B., Bennett, W.S., Strominger, J.L., and Wiley, D.C. (1987) Structure of the human class-I histocompatibility antigen, HLA-A2. *Nature 329*, 506-512.

Blaho, D. V., and Miranker, A. D. (2009) Delineating the conformational elements responsible for Cu2+-induced oligomerization of beta-2 microglobulin. *Biochemistry 48*, 6610-6617.

Bucciantini, M., Giannoni, E., Fabrizio, C., Baroni, F., Formigli, L., Zurdo, J., Taddei, N., Ramponi, G., Dobson, C.M., and Stefani, M. (2002) Inherent toxicity of aggregates implies a common mechanism for protein misfolding diseases. *Nature 416*, 507-511.

Bush, A. I. and Tanzi, R. E. (2002) The galvanization of beta-amyloid in Alzheimer's disease, *Proc. Natl. Acad. Sci. USA 99*, 7317-7319.

Calabrese, M.F., and Miranker, A.D. (2007) Formation of a stable oligomer of β-2 microglobulin requires only a transient encounter with Cu(II). *J. Mol. Biol. 367*, 1-7.

Calabrese, M. F., Eakin, C. M., Wang, J. M., and Miranker, A. D. (2008) A regulatable switch mediates self-association in an immunoglobulin fold. *Nature Struct. Mol. Biol. 15*, 965-971.

Calvete, J.J., Campanero-Rhodes, M.A., Raida, M., and Sanz, L. (1999) Characterization of the conformational and quaternary structure-dependent heparin-binding region of bovine seminal plasma protein PDC-109. *FEBS Letters 444*, 260-264.

Camerman, N., Camerman, A., and Sarkar, B. *(*1976*)* Molecular design to mimic the copper(II) transport site of human albumin. The crystal and molecular structure of copper(II) – glycylglycyl-L-histidine-*N*-methyl amide monoaquo complex. *Can. J. Chem. 54*, 1309–1316.

Carven, G.J., and Stern, L.J. (2005) Probing the ligand-induced conformational change in HLA-DR1 by selective chemical modification and mass spectrometric mapping. *Biochemistry 44*, 13625-13637.

Chen, T. and Weng, Z. (2003) ZDOCK: an initial-stage protein-docking algorithm. *Proteins 52*, 80-87.

Chen, T., Li, L., Weng, Z. (2002) Docking unbound proteins using shape complementary, desolvation, and electrostatics. *Proteins 47*, 281-294.

Cole, N.B., Murphy, D.D., Lebowitz, J., Di Noto, L., Levine, R.L., and Nussbaum, R.L. (2005) Metal-catalyzed oxidation of alpha-synuclein: helping to define the relationship between oligomers, protofibrils, and filaments. *J. Biol. Chem. 280*, 9678–9690.

Cornell, W., Cieplak, P., Bayly, C., Gould, I., Merz, K., Ferguson, D., Spellmeyer, D., Fox, T., Caldwell, J., and Kollma, P. (1995) A second generation force field for the simulation of proteins, nucleic acids, and organic molecules. *J. Am. Chem. Soc. 117*, 5179-5197.

Dage, J.L., Sun, H., and Halsall, H.B. (1998) Determination of diethylpyrocarbonate-modified amino acid residues in α_1-acid glycoprotein by high-performance liquid chromatography electrospray ionization-mass spectrometry and matrix-assisted laser desorption/ionization-time of flight mass spectrometry. *Anal. Biochem. 257*, 176-185.

Davis, D. P., Gallo, G., Vogen, S. M., Dul, J. L., Sciarretta, K. L., Kumar, A., Raffen, R., Stevens, F. J., and Argon, Y. (2001) Both the environment and somatic mutations govern the aggregation pathway of pathogenic immunoglobulin light chain, *J. Mol. Biol. 313*, 1021-1034.

Dobson, C.M. (1999) Protein misfolding, evolution and disease. *Trends Biochem. Sci. 24*, 329-330.

Eakin, C. M. and Miranker, A. D. (2005) From chance to frequent encounters: origins of β2-microglobulin fibrillogenesis. *Biochim. Biophys. Acta 1753*, 92-99.

Dong, F., and Zhou, H.-X. (2006) Electrostatic Contribution to the Binding Stability of Protein–Protein Complexes. *Proteins 65*, 87–102.

Eakin, C. M., Knight, J. D., Morgan, C. J., Gelfand, M. A., and Miranker, A. D. (2002) Formation of a copper specific binding site in non-native states of β-2-microglobulin, *Biochemistry 41*, 10646-10656.

Eakin, C.M., Attenello, F.J., Morgan, C.J., and Miranker, A.D. (2004) Oligomeric assembly of native-like precursors precedes amyloid formation by β-2 microglobulin. *Biochemistry 43*, 7808-7815.

Eakin, C.M., Berman, A.J., and Miranker, A.D. (2006) A native to amyloidogenic transition regulated by a backbone trigger. *Nature Struct. Mol. Biol. 13*, 202-208.

Ehrhardm, B., Misselwitz, R., Welfle, K., Hausdorf, G., Glaser, R.W., Schneider-Mergener, J., and Welfle, H. (1996) Chemical modification of recombinant HIV-1 capsid protein p24 leads to the release of a hidden epitope prior to changes of the overall folding of the protein. *Biochemistry 35*, 9097-9105.

Eichner, T., and Radford, S.E. (2009) A generic mechanism of β2-microglobulin amyloid assembly at neutral pH involving a specific proline switch. *J. Mol. Biol. 386,* 1312-1326.

Esposito, G., Michelutti, R., Verdone, G., Viglino, P., Hernandez, H., Robinson, C. V., Amoresano, A., Dal Piaz, F., Monti, M., Pucci, P., Mangione, P., Stoppini, M., Merlini, G., Ferri, G., and Bellotti, V. (2000) Removal of the N-terminal hexapeptide from human β2-microglobulin facilitates protein aggregation and fibril formation, *Protein Sci. 9*, 831-845.

Eyles, S.J., and Kaltashov, I.A. (2004) Methods to study protein dynamics and folding by mass spectrometry. *Methods 34*, 88-99.

Ferrao-Gonzales, A.D., Robbs, B.K., Moreau, V.H., Ferreira, A., Juliano, L., Valente, A.P., Almeida, F.C.L., Silva, J.L., and Foguel, D. (2005) Controlling beta-amyloid oligomerization by the use of naphthalene sulfonates – trapping low molecular weight oligomeric species. *J. Biol. Chem. 280*, 34747-34754.

Fliss, H., and Viswanatha, T. (1979) 2,3-Butanedione as a photosensitizing agent – Application to alpha-amino acids and alpha-chymotrypsin. *Can. J. Biochem. 57*, 1267-1272.

Floege, J., and Ehlerding, G. (1996) Beta-2-microglobulin associated amyloidosis. *Nephron 72*, 9-26.

Fraczkiewicz, R., and Braun, W. (1998) Exact and Efficient analytical calculation of the accessible surface areas and their gradients for macromolecules. *J. Comp. Chem. 19*, 319-333.

Friedhoff, P., von Bergen, M., Mandelkow, E.M., Davies, P., and Mandelkow, E. (1998) A nucleated assembly mechanism of Alzheimer paired helical filaments. *Proc. Natl. Acad. Sci. USA 95*, 15712-15717.

Gabant, G., Augier, J., and Armengaud, J. (2008) Assessment of solvent residues accessibility using three Sulfo-NHS-biotin reagents in parallel: application to footprint changes of a methyltransferase upon binding its substrate. *J. Mass Spectrom. 43*, 360–370.

Gao, J., Zhang, F., Zhang, L., Guo, Y., Ruan, K., Jiang, D., and Xu, C. (2007) Six specific lysine residues are crucial in maintaining the structure and function of soluble manganese stabilizing protein. *Acta Biochim. et Biophys. Sinica* **38**, 611-619.

Gao, Y., and Wang, Y. (2006) Site-selective modifications of arginine residues in human hemoglobin induced by methylglyoxal. *Biochemistry 45*, 15654-15660.

Giasson, B.I., Murray, I.V.J., Trojanowski, J.Q., and Lee, V.M.Y. (2001) A hydrophobic stretch of 12 amino acid residues in the middle of alpha-synuclein is essential for filament assembly. *J. Biol. Chem. 276*, 2380-2386.

Glocker, M.O., Borchers, C., Fiedler, W., Suckau, D., and Przybylski, M. (1994) Molecular characterization of surface topology in protein tertiary structures by amino-acylation and mass spectrometric peptide mapping. *Bioconjugate Chem. 5*, 583-590.

Glocker, M.O., Kalkum, M., Yamamoto, R., and Schreurs, J. (1996) Selective biochemical modification of functional residues in recombinant human macrophage colony-stimulating factor β (rhM-CSF β): Identification by mass spectrometry. *Biochemistry 35*, 14625-14633.

Guan, J.Q., and Chance, M.R. (2005) Structural proteomics of macromolecular assemblies using oxidative footprinting and mass spectrometry. *Trends Biochem. Sci. 30*, 583-592.

Hassani, O., Mansuelle, P., Cestele, S., Bourdeaux, M., Rochat, H., and Sampieri, F. (1999) Role of lysine and tryptophan residues in the biological activity of toxin VII (Ts gamma) from the scorpion *Tityus serrulatus*. *Eur. J. Biochem. 260*, 76-86.

Hochleitner, E.O., Borchers, C., Parker, C., Bienstock, R.J., and Tomer, K.B. (2000) Characterization of a discontinuous epitope of the human immunodeficiency virus (HIV) core protein p24 by epitope excision and differential chemical modification followed by mass spectrometric peptide mapping analysis. *Protein Sci. 9*, 487-496.

Holst, M. J., Baker, N. A., and Wang, F. (2000) Adaptive multilevel finite element solution of the Poisson-Boltzmann equation. I. Algorithms and examples. *J. Comput. Chem. 21*, 1319-1242.

Jeng, J., Lin, M.-F., Cheng, F.-Y., Yeh, C.-S., and Shiea, J. (2007) Using high-concentration trypsin-immobilized magnetic nanoparticles for rapid in situ protein digestion at elevated temperature. *Rapid Comm. Mass Spectrom. 21,* 3060-3068.

Jin, X.R., Abe, Y., Li, C.Y., and Hamasaki, N. (2003) Histidine-834 of human erythrocyte band 3 has an essential role in the conformational changes that occur during the band 3-mediated anion exchange. *Biochemistry 42,* 12927-12932.

Jobling, M. F., Huang, X., Stewart, L. R., Barnham, K. J., Curtain, C., Volitakis, I., Perugini, M., White, A. R., Cherny, R. A., Masters, C. L., Barrow, C. J., Collins, S. J., Bush, A. I., and Cappai, R. (2001) Copper and zinc binding modulates the aggregation and neurotoxic properties of the prion peptide PrP106-126. *Biochemistry 40,* 8073-8084.

Kad, N.M., Myers, S.L., Smith, D.P., Smith, D.A., Radford, S.E., and Thomson, N.H. (2003) Hierarchical assembly of β_2-microglobulin amyloid in vitro revealed by atomic force microscopy. *J. Mol. Biol. 330,* 785-797.

Kalkum, M., Przybylski, M., and Glocker, M.O. (1998) Structure characterization of functional histidine residues and carbethoxylated derivatives in peptides and proteins by mass spectrometry. *Bioconjugate Chem. 9,* 226-235.

Kaltashov, I. A., and Mohimen, A. (2005) Estimates of protein surface areas in solution by electrospray ionization mass spectrometry. *Anal. Chem. 77,* 5370-5379.

Kammerer, R.A., Kostrewa, D., Zurdo, J., Detken, A., Garcia-Echeverria, C., Green, J.D., Muller, S.A., Meier, B.H., Winkler, F.K., Dobson, C.M., and Steinmetz, M.O. (2004) Exploring amyloid formation by a de novo design. *Proc. Natl. Acad. Sci. USA 101,* 4435-4440.

Keating, M.J. (1999) Chronic lymphocytic leukemia. *Semin. Oncol. 26,* 107-114.

Kelly, J.W. (1998) The alternative conformations of amyloidogenic proteins and their multistep assembly pathways. *Curr. Opin. Biol. 8,* 101-106.

Keskin, O.; Gursoy, A.; Ma, B. and Nussinov, R. (2008) Principles of protein-protein interactions: What are the preferred ways for proteins to interact? *Chem. Rev. 108,* 1225-1244.

Khan, A. R., Baker, B. M., Ghosh, P., Biddison, W. E., and Wiley, D. C. (2000) The structure and stability of an HLA-A*0201/octameric tax peptide complex with an empty conserved peptide-N-terminal binding site. *J. Immunol. 164,* 6398-6405.

Kihara, M., Chatani, E., Iwata, K., Yamamoto, K., Matsuura, T., Nakagawa, A., Naiki, H., and Goto, Y. (2006) Conformation of amyloid fibrils of β2-microglobulin probed by tryptophan mutagenesis. *J. Biol. Chem. 281,* 31061-31069.

Kim, W., and Hecht, M.H. (2006) Generic hydrophobic residues are sufficient to promote aggregation of the Alzheimer's A beta 42 peptide. *Proc. Natl. Acad. Sci. USA 103*, 15824-15829.

Kim, J.Y., Kim, K.W., Kwon, H.J., Lee, D.W., and Yoo, J.S. (2002) Probing lysine acetylation with a modification-specific marker ion using high-performance liquid chromatography/electrospray-mass spectrometry with collision-induced dissociation. *Anal. Chem. 74*, 5443-5449

Knock, S.L., Miller, B.T., Blankenship, J.E., Nagle, G.T., Smith, J.S., and Kurosky, A. (1991) N-acylation of aplysia egg-laying hormone with biotin. *J. Biol. Chem. 266*, 24413-24419.

Kodre, A., Arcon, I., and Gomilsek, J.P. (2004) X-ray absorption spectroscopy and related techniques. *Acta Chim. Slovenica 51*, 1-10.

Komives, E.A. (2005) Protein-protein interaction dynamics by amide H/H-2 exchange mass spectrometry. *Int. J. Mass Spec. 24*0, 285-290.

Kundrotas, P. J., and Alexov, E. (2006) Electrostatic Properties of Protein-Protein Complexes. *Biophys. J. 91*, 1724–1736.

Leitner, A., Amon, S., Rizzi, A., and Lindner, W. (2007) Use of the arginine-specific butanedione/phenylboronic acid tag for analysis of peptides and protein digests using matrix-assisted laser desorption/ionization mass spectrometry. *Rapid Comm. Mass Spectrom. 21*, 1321-1330.

Leitner, A., and Lindner, W. (2005) Functional probing of arginine residues in proteins using mass spectrometry and an arginine-specific covalent tagging concept. *Anal. Chem. 77*, 4481-4488.

Lim, J. and Vachet, R. W. (2004) Using mass spectrometry to study copper-protein binding under native and non-native conditions: β-2-microglobulin, *Anal. Chem. 76*, 3498-3504.

Lindahl, E., Hess, B., and van der Spoel, D. (2001) Gromacs 3.0: A package for molecular simulation and trajectory analysis. *J. Mol. Mod. 7*, 306-317.

Liu, Y.Y., Kvaratskhelia, M., Hess, S., Qu, Y.X., and Zou, Y. (2005) Modulation of replication protein A function by its hyperphosphorylation-induced conformational change involving DNA binding domain B. *J. Biol. Chem. 280*, 32775-32783.

López de la Paz, M., Goldie, K., Zurdo, J., Lacroix, E., Dobson, C.M., Hoenger, A., and Serrano, L. (2002) *De novo* designed peptide-based amyloid fibrils. *Proc. Natl. Acad. Sci. USA 99*, 16052-16057.

Lundblad, R.L., and Noyes, C.M. (1984) Chemical reagents for protein modification. CRC Press, Boca Raton Fl. Ch. 1, 105-126.

Mahoney, M. Jorgensen, A. (2000) A five-site model for liquid water and the reproduction of the density anomaly by rigid, nonpolarizable potential functions. *J. Chem. Phys. 112*, 8911-8922.

Mailfait, S., Belaiche, D., Kouach, M., Dallery, N., Chavette, P., Formstecher, P., and Sablonnière, B. (2000) Critical role of tyrosine 277 in the ligand-binding and transactivating properties of retinoic acid receptor alpha. *Biochemistry 39*, 2183-2192.

Malaguarnera, M., Restuccia, S., Di Fazio, I., Zoccolo, A.M., Trovato, B.A., and Pistone, G. (1997) Serum beta-2-microglobulin in chronic hepatitis C. *Dig. Dis. Sci. 42*, 762-766.

Marie, G., Serani, L., Laprévote, O., Cahuzac, B., Guittet, E., and Felenbok, B. (2001) Differential chemical labeling of the AlcR DNA-binding domain from Aspergillas nidulans versus its complex with a 16-mer DNA target: Identification of an essential tryptophan involved in the recognition and the interaction with the nucleic acid. *Protein Sci. 10*, 99-107.

Massolini, G., and Calleri, E. (2005) An improved clean sonoreactor-based method for protein identification by mass spectrometry-based techniques. *J. Sep. Sci. 28*, 7-21.

McLafferty, F.W., Breuker, K., Jin, M., Han, X.M., Infusini, G., Jiang, H., Kong, X.L., and Begley, T.P. (2007) Top-down MS, a powerful complement to the high capabilities of proteolysis proteomics. *FEBS J. 274*, 6256-6268.

McParland, V.J., Kad, N.M., Kalverda, A.P., Brown, A., Kirwin-Jones, P., Hunter, M.G., Sunde, M., and Radford, S.E. (2000) Partially unfolded states of β_2-microglobulin and amyloid formation in vitro. *Biochemistry 39*, 8735-8746.

McParland, V.J., Kalverda, A.P., Homans, S.W., and Radford, S.E. (2002) Structural properties of an amyloid precursor of β_2-microglobulin. *Nature Struct. Biol. 9*, 326-331.

Melchoir, W.B. Jr, and Fahrney, D. (1970) Ethoxyformylation of proteins. Reaction of ethoxyformic anhydride with α-chymotrypsin, pepsin, and pancreatic ribonuclease at pH 4. *Biochemistry 9*, 251-258.

Mendes Sousa, M., Cardoso, I., Fernandes, R., Guimaraes, A., and Saraiva, M.J. (2001) Deposition of transthyretin in early stages of familial amyloidotic polyneuropathy. Evidence for toxicity of nonfibrillar aggregates. *Am. J. Pathol. 159*, 1993-2000.

Mendoza, V.L., and Vachet, R.W. (2008) Protein surface mapping using diethylpyrocarbonate with mass spectrometric detection. *Anal. Chem. 80*, 2895-2904.

Mendoza, V. L., and Vachet, R. W. (2009) Probing protein structure by amino acid-specific covalent labeling and mass spectrometry. *Mass Spec. Rev. 28*, 785-815.

Mendoza, V.L., Antwi, K., Baron-Rodriguez, M.A., Blanco, C., and Vachet, R.W. (2009) Structure of the preamyloid dimer of beta-2-microglobulin from covalent labeling and mass spectrometry. *Biochemistry 49*, 1522-1532.

Meng, F.Y., Forbes, A.J., Miller, L.M., and Kelleher, N.L. (2005) Detection and localization of protein modifications by high resolution tandem mass spectrometry. *Mass Spectrom. Rev. 24*, 126-134.

Miles, E.W. (1977) Modification of histidyl residues in proteins by diethylpyrocarbonate. *Method Enzymol. 47*, 431-442.

Morgan, C.J., Gelfans, M., Atreya, C., and Miranker, A.D. (2001) Kidney dialysis-associated amyloidosis: a molecular role for copper in fiber formation. *J. Mol. Biol. 309*, 339-345.

Musafia, B., Buchner, V., and Arad, D. (1995) Complex salt bridges in proteins: Statistical analysis of structure and function. *J Mol. Biol. 254*, 761-770.

Naiki, H., Hasegawa, K., Yamaguchi, I., Nakamura, H., Gejyo, F., and Nakakuki, K. (1998) Apolipoprotein E and antioxidants have different mechanisms of inhibiting Alzheimer's beta-amyloid fibril formation in vitro. *Biochemistry 37*, 17882-17889.

Narindrasorasak, S., Kulkarni, P., Deschamps, P., She, Y. M. and Sarkar, B. (2007) Characterization and copper binding properties of human COMMD1 (MURR1), *Biochemistry 46*, 3116-3128.

Novak, P., Kruppa, G.H., Young, M.M., and Schoeniger, J. (2004) A top-down method for the determination of residue-specific solvent accessibility in proteins. *J. Mass Spectrom. 39*, 322-328.

Nuss, J.E., Sweeney, D.J., and Alter, G.M. (2006) Reactivity-based analysis of domain structures in native replication protein A. *Biochemistry 45*, 9804-9818.

Ohhashi, Y., Kihara, M., Naiki, H., and Goto, Y. (2005) Ultrasonication-induced amyloid fibril formation of β_2-microglobulin, *J. Biol. Chem. 280*, 32843-32848.

Okon, M., Bray, P., and Vucelic, D. (1992) H^1-NMR assignments and secondary structure of human beta-2-microglobulin in solution. *Biochemistry 31*, 8906−8915.

Perkins, C.M., Roes, N.J., Weinstein, B., Stenkamp, R.E., Jensen, L.H., and Pickart, L. *(*1984*)* The structure of a copper complex of the growth-factor glycyl-L-histidyl-L-lysine at 1.1 A resolution. *Inorg. Chim. Acta* 82, 93–99.

Pramanik, B.M., Mirza, U.A., Ing, Y.-H., Liu, Y.-H., Bartner, P.L., Weber, P.C., and Bose, A.K. (2002) Microwave-enhanced enzyme reaction for protein mapping by mass spectrometry: A new approach to protein digestion in minutes. *Protein. Sci. 11*, 2676-2687.

Qin, K., Yang, Y., Mastrangelo, P., and Westaway, D. (2002) Mapping Cu(II) binding sites in prion protein by diethyl pyrocarbonate modification and matrix-assisted laser desorption ionization-Time of Flight (MALDI-TOF) mass spectrometric footprinting. *J. Biol. Chem. 277*, 1981-1990.

Relini, A., Canale, C., De Stefano, S., Rolandi, R., Giorgetti, S., Stoppini, M., Rossi, A., Fogolari, F., Corazza, A., Esposito, G., Gliozzi, A., and Bellotti, V. (2006) Collagen plays an active role in the aggregation of β_2-microglobulin under physiopathological conditions of dialysis-related amyloidosis, *J. Biol. Chem. 281*, 16521-16529.

Rial-Otero, R., Carreira, R.J., Cordeiro, F.M., Moro, A.J., Santos, H.M., Vale, G., Moura, I., and Capelo, J.L. (2007) Ultrasonic assisted protein enzymatic digestion for fast protein identification by matrix-assisted laser desorption/ionization time-of-flight mass spectrometry Sonoreactor versus ultrasonic probe. *J. Chromatogr. A* 1166, 101-107.

Riordan, J.F. (1973) Functional arginyl residues in carboxypeptidase-A – Modification with butanedione. *Biochemistry 12*, 3915-3923.

Riordan, J.F. (1979) Arginyl residues and anion binding-sites in proteins. *Mol. Cell Biochem. 26*, 71–92.

Risal, D., Gourinath, S., Himmel, D. M., Szent-Gyorgyi, A. G., and Cohen, C. (2004) Myosin subfragment 1 structures reveal a partially bound nucleotide and a complex salt bridge that helps couple nucleotide and actin binding. *Proc. Natl. Acad. Sci. USA 101*, 8930-8935.

Rochet, J.C., and Lansbury, P.T. Jr. (2000) Amyloid fibrillogenesis: themes and variations. *Curr. Opin. Struct. Biol. 10*, 60-68.

Royer, C.A., Mann, C.J., and Matthews, C.R. (1993) Resolution of the fluorescence equilibrium unfolding profile of *trp* aporepressor using single tryptophan mutants. *Protein Sci. 2*, 1844-1852.

Santos, H.M., Mota, C., Lodeiro, C., Moura, I., Isaac, I., and Capelo, J.L. (2008) An improved clean sonoreactor-based method for protein identification by mass spectrometry-based techniques. *Talanta 77*, 870-875.

Šantrůček, J., Strohalm, M., Kadlčík, V., Hynek, R., and Kodíček, M. (2004) Tyrosine residues modification studied by MALDI-TOF mass spectrometry. *Biochem. Biophys. Res. Comm. 323*, 1151-1156.

Schepens, I., Ruelland, E., Miginiac-Maslow, M., Marechal, P., and Decottignies, P. (2000) The role of active site arginines of Sorghum NADP-malate dehydrogenase in thioredoxin-dependent activation and activity. *J. Biol. Chem. 275*, 35792-35798.

Scholten, A., Visser, N.F.C., van den Heuvel, R.H.H., and Heck, A.J.R. (2006) Analysis of protein-protein interaction surfaces using a combination of efficient lysine acetylation and nanoLC-MALDI-MS/MS applied to the E9:Im9 bacteriotoxin-immunity protein complex. *J. Am. Soc. Mass Spectrom. 17*, 983-994.

Sheinerman, F. B.; Norel, R. and Honig, B. (2000) Electrostatic aspects of protein-protein interactions. *Curr. Opin. Struct. Biol. 10*, 153-159.

Smith, A.M., Jahn, T.R., Ashcroft, A.E., and Radford, S.E. (2006) Direct observation of oligomeric species formed in the early stages of amyloid fibril formation using electrospray ionization mass spectrometry. *J. Mol. Biol. 364*, 9-19.

Sorin, E. J. and Pande, V. S. (2005) Exploring the helix-coil transition via all-atom equilibrium ensemble simulations. *Biophys. J. 88*, 2472-2493.

Sperry, J.B., Shi, X., Rempel, D.L., Nishimura, Y., Akashi, S., and Gross, M.L. (2008) A mass spectrometric approach to the study of DNA-binding proteins: interaction of human TRF2 with telomeric DNA. *Biochemistry 47*, 1797-1807.

Stefani, M. (2004) Protein misfolding and aggregation: new examples in medicine and biology of the dark side of the protein world. *Biochim. Biophys. Acta 1739*, 5-25.

Steiner, R.F., Albaugh, S., Fenselau, C., Murphy, C., and Vestling, M. (1991) A Mass-spectrometry method for mapping the interface topography of interacting proteins, illustrated by the melittin-calmodulin system. *Anal. Biochem. 196*, 120-125.

Strohalm, M., Šantrůček, J., Hynek, R., and Kodíček, M. (2004) Analysis of tryptophan surface accessibility in proteins by MALDI-TOF mass spectrometry. *Biochem. Biophys. Res. Comm. 323*, 1134-1138.

Suckau, D., Mak, M., and Przybylski, M. (1992) Protein surface topology-probing by selective chemical modification and mass spectrometric peptide mapping. *Proc. Natl. Acad. Sci. USA 89*, 5630-5634.

Sun, W., Gao, S., Wang, L., Chen, Y., Wu, S., Wang, X., Zheng, D., and Gao, Y. (2005) Microwave-assisted protein preparation and enzymatic digestion in proteomics. *Mol. Cell. Proteomics 5*, 769-776.

Sunde, M., Serpell, L.C., Bartlam, M., Fraser, P.E., Pepys, M.B., and Blake, C.C. (1997) Common core structure of amyloid fibrils by synchroton X-ray diffraction. *J. Mol. Biol. 273*, 729-739.

Takamoto, K., and Chance, M.R. (2006) Radiolytic protein footprinting with mass Spectrometry to probe the structure of macromolecular complexes. *Annu. Rev. Biophys. Biomol. Struct. 35*, 251-276.

Tong, X., Wren, J.C., and Konermann, L. (2007) Effects of protein concentration on the extent of gamma-ray-mediated oxidative Labeling studied by electrospray mass spectrometry. *Anal. Chem. 79*, 6376-6382.

Trinh, C.H., Smith, D.P., Kalverda, A.P., Phillips, S.E.V., and Radford, S.E. (2002) Crystal structure of monomeric human beta-2-microglobulin reveals clues to its amyloidogenic properties. *Proc. Natl. Acad. Sci.USA 99*, 9771-9776.

Trinquier, G., and Sanejouand, Y.H. (1998) Which effective property of amino acids is best preserved by the genetic code? *Protein Eng. 11*, 153-169.

Tsubaki, M., Kobayashi, K., Ichise, T., Takeuchi, F., and Tagawa, S. (2000) Diethyl pyrocarbonate modification abolishes fast electron accepting ability of cytochrome b(561) from ascorbate but does not influence electron donation to monodehydroascorbate radical: Identification of the modification sites by mass spectrometric analysis. *Biochemistry 39*, 3276-3284.

Tycko, R. (2004) Progress towards a molecular-level structural understanding of amyloid fibrils. *Curr. Opin. Struct. Biol. 14*, 96-103.

Uversky, V. N., Li, J., and Fink, A. L. (2001) Metal-triggered structural transformations, aggregation, and fibrillation of human alpha-synuclein. a possible molecular link between Parkinson's disease and heavy metal exposure, *J. Biol. Chem. 276*, 44284-44296.

van der Spoel, D., Lindahl, E., Hess, B., Groenhof, G., Mark, A., and Berendsen, H. (2005) Gromacs: Fast, flexible, and free. *J. Comp. Chem. 26*, 1701-1719.

Van Horn, J.D., Bulaj, G., Goldenberg, D.P., and Burrows, C. J. J. (2003) The Cys-Xaa-His metal-binding motif: {N} versus {S} coordination and nickel-mediated formation of cysteinyl sulfinic acid. *Biol. Inorg. Chem. 8*, 601– 610

Verdone, G., Corazza, A., Viglino, P., Pettirossi, F., Giorgetti, S., Mangione, P., Andreola, A., Stoppini, M., Bellotti, V., and Esposito, G. (2002) The solution structure of human β2-microglobulin reveals the prodromes of its amyloid transition. *Protein Sci. 11*, 487-499.

Viles, J.H., Cohen, F.E., Prusiner, S.B., Goodin, D.B., Wright, P.E., and Dyson, H.J. (1999) Copper binding to the prion protein: Structural implications of four identical cooperative binding sites. *Proc. Natl. Acad. Sci. USA 96*, 2042-2047.

Villanueva, J., Hoshino, M., Katou, H., Kardos, J., Hasegawa, K., Naiki, H., and Goto, Y. (2004) Increase in the conformational flexibility of β_2-microglobulin upon copper binding: a possible role for copper in dialysis-related amyloidosis. *Protein Sci. 13*, 797-809.

Wadsworth, J. D., Hill, A. F., Joiner, S., Jackson, G. S., Clarke, A. R., and Collinge, J. (1999) Strain-specific prion-protein conformation determined by metal ions. *Nature Cell Biol. 1*, 55-59.

Wales, T.E., and Engen, J.R. (2006) Hydrogen exchange mass spectrometry for the analysis of protein dynamics. *Mass Spectrom. Rev.25*, 158-170.

Walsh, D.M., Lomakin, A., Benedek, G.B., Condron, M.M., and Teplow, D.B. (1997) Amyloid beta-protein fibrillogenesis – detection of a protofibrillar intermediate. *J. Biol. Chem. 272*, 22364-22372.

Wang, X., Kim, S.H., Ablonczy, Z., Crouch, R.K., and Knapp, D.R. (2004) Probing rhodopsin-transducin interactions by surface modification and mass spectrometry. *Biochemistry 43*, 11153-11162.

Wen, J.Z., Zhang, H., Gross, M.L., and Blankenship, R.E. (2009) Membrane orientation of the FMO antenna protein from *Chlorobaculum tepidum* as determined by mass spectrometry-based footprinting. *Proc. Natl. Acad. Sci. USA 106*, 6134-6139.

Wuerges, J., Lee, J.W., Yim, Y.I., Yim, H.S., Kang, S.O., and Carugo, K.D. (2004) Crystal structure of nickel-containing superoxide dismutase reveals another type of active site. *Proc. Natl. Acad. Sci. USA 101*, 8569– 8574.

Xia, W., Li, H.Y., Sze, K.H., and Sun, H.Z. (2009) Structure of a nickel chaperone, HypA, from *Helicobacter pylori* reveals two distinct metal binding sites. *J. Am. Chem. Soc. 131*, 10031-10040.

Xu, D.; Lin, S. L. and Nussinov, R. (1997) Protein binding *versus* protein folding: The role of hydrophilic bridges in protein associations. *J. Mol. Biol. 265*, 68-84.

Xu, G. Z., Liu, R. T., Zak, O., Aisen, P., and Chance, M. R. (2005). Structural allostery and binding of the transferrin·receptor complex. *Mol. Cel.. Proteomics 4*, 1959-1967.

Yamaguchi, K., Katou, H., Hoshino, M., Hasegawa, K., Naiki, H., and Goto, Y. (2004) Core and heterogeneity of β_2-microglobulin amyloid fibrils as revealed by H/D exchange. *J. Mol. Biol. 338*, 559-571.

Yem, A.W., Epps, D.E., Mathews, W.R., Guido, D.M., Richard, K.A., Staite, N.D., and Deibel, M.R. (1992) Site-specific chemical modification of interleukin-1 beta by acrylodan at cysteine-8 and lysine-103. *J. Biol. Chem. 267*, 3122-3128.

Zatta, P., Drago, D., Bolognin, S., and Sensi, S.L. (2009) Alzheimer's disease, metal ions and metal homeostatic therapy. *Trends Pharmacol. Sci. 30*, 346-355.

Zerovnik, E. (2002) Amyloid-fibril formation: proposed mechanisms and relevance to conformational disease. *Eur. J. Biochem. 269*, 3361-3371.

Zhai, H.L., Dorrestein, P.C., Chatterjee, A., Begley, T.P., and McLafferty, F.W. (2005) Simultaneous kinetic characterization of multiple protein forms by top down mass spectrometry. *J. Am. Soc. Mass Spectrom. 16*, 1052-1059.